

AD-A207 848

DTIC

②

DOCUMENT PAGE

Form Approved  
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED			b. RESTRICTIVE MARKINGS FILE		
2a. SECURITY CLASSIFICATION AUTHORITY E			3. DISTRIBUTION/AVAILABILITY OF REPORT APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED		
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFOSR/ONR TECH REPORT 89-1		
5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR-89-0628			6a. NAME OF PERFORMING ORGANIZATION HARVARD UNIVERSITY		
6b. OFFICE SYMBOL (If applicable)			7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research		
6c. ADDRESS (City, State, and ZIP Code) 1236 William James Hall 33 Kirkland Street Cambridge, MA 02138-2044			7b. ADDRESS (City, State, and ZIP Code) Building 410 Bolling Air Force Base, D.C. 20332-6448		
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR			8b. OFFICE SYMBOL (If applicable) NL		
9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR 88-0012			10. SOURCE OF FUNDING NUMBERS		
8c. ADDRESS (City, State, and ZIP Code) Building 410 Bolling Air Force Base, D.C. 20332-6448			PROGRAM ELEMENT NO. 61102F		
			PROJECT NO. 2313		
			TASK NO. A4		
			WORK UNIT ACCESSION NO.		
11. TITLE (Include Security Classification) Components of High-level Vision: A Cognitive Neuroscience Analysis + Accounts of Neurological Syndromes					
12. PERSONAL AUTHOR(S) Stephen M. Kosslyn					
13a. TYPE OF REPORT Annual Tech. Report		13b. TIME COVERED FROM 12/87 TO 2/89		14. DATE OF REPORT (Year, Month, Day) 1989 Feb. 28	
				15. PAGE COUNT 90	
16. SUPPLEMENTARY NOTATION					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP	Neuropsychology, vision, computational models, spatial orientation. (AW)		
05	09				
19. ABSTRACT (Continue on reverse if necessary and identify by block number) High-level visual processes make use of stored information, and are invoked during object identification, navigation, tracking, and visual mental imagery. The present work has revolved around a theory of the component processing subsystems used in high-level vision. This theory was developed by considering neuroanatomical, neurophysiological, and computational constraints. The theory has led to three kinds of empirical work: First, specific claims associated with individual processing subsystems have been tested. For example, the analysis of the representation of spatial relations led to the prediction that two subsystems are used to encode this information, and a set of experiments was conducted that provided support for this distinction. Second, predictions from the theory as a whole have been formulated, and some of these predictions are now being tested. And third, the subsystems have been implemented in a running computer simulation model, which has been used to generate predictions about specific neurological syndromes. The model can be damaged in a variety of ways, and its performance on a set of tasks then observed. The experiments conducted to date and predictions from the computer model are summarized in this report. In addition, the most common dysfunctions of vision following brain damage are reviewed, and accounts are offered by reference to the simulation model. Keywords:					
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED		
22a. NAME OF RESPONSIBLE INDIVIDUAL Alfred R. Fregly			22b. TELEPHONE (Include Area Code) (202) 767-5021		22c. OFFICE SYMBOL NL

DD Form 1473, JUN 86

Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

89 5 15

122

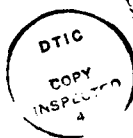
UNCLASSIFIED

0 MAR 1989

## Abstract

High-level visual processes make use of stored information, and are invoked during object identification, navigation, tracking, and visual mental imagery. The present work has revolved around a theory of the component "processing subsystems" used in high-level vision. This theory was developed by considering neuroanatomical, neurophysiological, and computational constraints. The theory has led to three kinds of empirical work: First, specific claims associated with individual processing subsystems have been tested. For example, the analysis of the representation of spatial relations led to the prediction that two subsystems are used to encode this information, and a set of experiments was conducted that provided support for this distinction. Second, predictions from the theory as a whole have been formulated, and some of these predictions are now being tested. And third, the subsystems have been implemented in a running computer simulation model, which has been used to generate predictions about specific neurological syndromes. The model can be damaged in a variety of ways, and its performance on a set of tasks then observed. The experiments conducted to date and predictions from the computer model are summarized in this report. In addition, the most common dysfunctions of vision following brain damage are reviewed, and accounts are offered by reference to the simulation model.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input checked="" type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



**APOSR - TR. 89 - 0628**

**COMPONENTS OF HIGH-LEVEL VISION:  
A COGNITIVE NEUROSCIENCE ANALYSIS  
AND ACCOUNTS OF NEUROLOGICAL SYNDROMES**

*In recent years there has been tremendous excitement about neural network models of cognitive processes. Not only can these networks learn, but they sometimes provide insight into otherwise puzzling phenomena (Rumelhart and McClelland, 1986a). However, very little effort has been expended on determining which tasks should be solved by a network. Indeed, in most cases the decision about what task to model is made entirely on intuitive grounds, sometimes with less than salubrious consequences. For example, Rumelhart and McClelland (1986b) assumed that a network should be dedicated to computing the past tense of verbs, and developed a network that performed this task reasonably well. However, Pinker and Prince (1988) showed that linguistic and other criteria render implausible the assumption that a single mechanism computes only the past tense, and it appears unlikely that such a network will ever perform the task correctly. One goal of this article is to explicate a set of information-processing tasks that might be carried out by distinct neural networks.*

*A second goal of this article grows out of the first. The idea that mental activity is carried out by a collection of distinct components is hardly new; indeed, the founders of modern neurology were well aware of this idea (e.g., Jackson, 1864), as were many of their predecessors. However, there often has been the assumption that the functional components will bear a direct relation to components of observable behavior. An extreme example is evident in the claims of the phrenologists, who assumed that distinct parts of the brain were responsible for charity, hopefulness, and love, among other behaviors. Marr (1982) offered a different approach to characterizing components, based on the notion that brain function can be understood as computation. In the Marrian approach, one motivates a componential breakdown by considering what is necessary for a mechanism to produce a given behavior in specific circumstances. Kosslyn (1987) used this approach to begin developing the foundations for a neuropsychologically plausible theory of high-level vision, and the present article continues in this direction, now providing a relatively detailed characterization of the structure of the system.*

*In particular, in this article we consider a variety of neurological syndromes that affect vision, and show that the range of possible causes for these deficits is far richer than is currently appreciated in the neuropsychological literature. We not only develop an explicit theory of the underlying processing components, but embody this theory in a running computer simulation model that makes explicit predictions. Thus, our aim is to replace the current taxonomies of syndromes with a different sort of account, based on disruptions of the underlying information processing.*

## I. SUBSYSTEMS OF HIGH-LEVEL VISION

The present article focuses on the information processing underlying high-level vision. The idea that visual processing can be divided into "high" and "low" levels stems in part from the recent discovery that a large number of areas of the primate brain are used to process visual information (e.g., see Van Essen, 1985). These areas have different functional properties, and apparently are involved in carrying out different kinds of computations. Low-level visual processing is driven by sensory input, and is concerned with using such input to find edges, grow regions of homogeneous texture or color, establish depth, and other tasks that will help one to segregate figure from ground. These areas typically are topographically organized, with adjacent parts of an image being processed by adjacent local patches. In contrast, high-level visual processing involves the use of previously stored information, and is concerned with using such information to identify objects, navigate, and form and use mental images. The areas that carry out such processing often are not topographically organized, and are physically farther removed from the locus of visual input from the eyes.

The theory of high-level vision developed in this article is unlike previous ones in a number of respects. Not only does it focus on the nature of the component processing units that underlie high-level vision, assuming that each one corresponds to a neural network or set of neural networks, but the theory is motivated in part by neurophysiology and neuroanatomy. In addition, the same mechanisms are intended to account for behavioral findings about visual perception and visual mental imagery (although we will not consider imagery in detail in the present article). Finally, the theory is intended to account not only for high-level visual processing in normal people, but also for the nature of behavioral dysfunctions following brain damage.

### Motivation for the Theory

The theory was developed in light of three kinds of considerations. First, we considered the abilities of the intact high-level visual system; any theory of visual dysfunction must be cast within a conception of the normal system. Second, we examined relevant neuroanatomical and neurophysiological findings, primarily from nonhuman primates. Third, in light of the foregoing we performed information-processing analyses, which led us to hypothesize distinct processing subsystems that – working together – are in principle capable of producing the observed functional properties of the system. The key points of each of the first two sources of motivation are summarized below; the information processing analyses will be presented as we formulate hypotheses of specific subsystems.

### Functions of the intact system

Any model of the effects of brain damage on behavior rests on assumptions, explicit or implicit, about the operation of the normal system. Without a characterization of the functioning of the intact system, it is very difficult to explain its dysfunctions. Thus, it behooves us to pause and briefly consider the essential characteristics of the functions of the system we wish to understand. In this article we focus on object recognition and identification. By "recognition" we mean achieving a sense of

familiarity by matching the visual representation against a previously stored one; by "identification" we mean not only realizing that a stimulus is familiar, but also having access to information associated with the object, such as its name, a description of its properties, and so on. The fundamental function of the brain systems underlying object identification is to know more about a stimulus than is apparent in the immediate input. For example, upon being shown an apple, one knows that it has seeds, even though they are not visible. This is done, of course, by using the stimulus to activate the appropriate information previously stored in memory. These functions are difficult to understand in part because they are robust under three classes of situations:

#### *Viewpoint independence*

We typically can identify objects when they subtend different visual angles, either because they are at different distances or they are of different sizes. We also typically can identify objects when they are viewed from novel vantage points or are misoriented. Furthermore, we can identify objects when they appear in different parts of the visual field.

#### *Shape variations*

We typically can identify objects when their shapes do not exactly match the shapes of previously seen objects. This occurs when an object's parts vary in shape (such as occurs with chairs, which can have different shaped arms, legs, backs and so on) or when objects have or do not have optional parts (such as arms for a chair). We also typically can identify objects when the spatial relations among the parts varies, such as occurs when a person is standing, squatting, sitting, and so forth. It is these abilities that rule out simple template theories of shape recognition and identification (Neisser, 1967).

#### *Impoverished input*

We typically can identify objects even when they are seen a part at a time, as occurs when they subtend large visual angles and multiple eye movements are required to encode the shape with high resolution. We typically can identify objects when they are partially occluded, and when they are partially degraded in other ways (e.g., seen in poor lighting or through a heavy fog).

Each of these general properties can be characterized more precisely (by observing the precise conditions under which correct performance begins to degrade, the processing time under different conditions, and so on). However, even at the most coarse level described here, many existing theories are in principle incapable of accounting for one or more of these properties, and hence can be eliminated. The challenge is to formulate a theory that posits mechanisms that are in principle capable of allowing the system as a whole to function as here described.<sup>1</sup>

#### Primary neuroanatomical and neurological constraints

Because the theory is a theory of how information is processed by the brain, we take seriously the known neuroanatomical and neurophysiological constraints. If one were intent on reading the literature on the brain in search of constraints on theories of visual information processing, it would

seem a hopeless task. There simply is too much known. However, if one has specific issues in mind, and searches only for information relevant to these particular concerns, much of use can be discovered rather easily. In developing the theory, four classes of constraints proved particularly useful, as is briefly summarized below.

#### *Retinotopic maps*

There are now approximately 30 distinct areas in the brain concerned with processing visual information (Van Essen, 1985; Van Essen, personal communication). Several features of these areas are particularly pertinent to the present concerns. Perhaps the most fundamental is the fact that some 10 of these areas preserve the local geometry of the retina (with magnification factors and other distortions; see Van Essen, 1985). That is, the image projected on the back of the retina is physically laid out on the back of the brain in multiple places. Tootell, Silverman, Switkes, and De Valois (1982) provided a particularly dramatic demonstration of this by having a monkey stare at a pattern after being injected with 2-deoxyglucose, a radioactive metabolic marker. The marker allowed them to see which groups of cells were most active when the animal was seeing the pattern. Sure enough, a picture of the pattern was literally projected onto the back of the brain (magnified at its center, in accordance with the greater representation given to the foveal areas of the retina), and could be "developed" and easily seen spread out on the largest visual area (VI).

#### *Two cortical visual pathways*

Perhaps the most striking neurological constraint is the evidence that *object properties* (such as shape and color) and *spatial properties* are processed in separate neural systems (for summaries, see Maunsell and Newsome, 1987; Ungerleider and Mishkin, 1982). The spatial properties pathway runs from the occipital lobe up to the parietal lobes, and has been called the "dorsal system;" the object properties pathway leads from the occipital lobe down to the inferior temporal lobe, and has been called the "ventral system." The dorsal pathway appears to receive input primarily from the low-level magnocellular pathways (which originate with large ganglion cells in the retina), whereas the ventral pathway appears to receive input primarily from the low-level parvocellular pathway (which originate with smaller ganglion cells in the retina), as characterized by Hubel and Livingstone (1987) and Livingstone and Hubel (1987a, 1987b). The magno pathway leads to the lateral geniculate nucleus (LGN), and then to layer 4B of area V1, and from there to the thick stripes of area V2. This pathway deals with movement and is very sensitive to binocular disparity and intensity differences. In contrast, the parvocellular ganglion cells project to different layers of the LGN than the magnocellular ganglion cells. These layers then project to the blobs and interblobs in area V1, which in turn project to the thin stripes and pale stripes (also known as "interstripes") in V2.

The division of the parvo pathway into two streams at the level of V1 and V2 corresponds to different functional properties. According to DeYeu and Van Essen (1988), cells in the blob-thin stripe stream appear to be sensitive to color and little else, whereas cells in the interblob-pale stripe stream

are sensitive to color, orientation, and binocular disparity. It is probably important to note that cells in at least this part of the parvo pathway respond to multiple stimulus dimensions.

The claim that the dorsal and ventral pathways correspond to distinct cortical visual systems, dealing (roughly speaking) with "what" and "where," is based on studies of the brain itself and studies of behavioral dysfunction following different sorts of brain damage (for a review, see Ungerleider and Mishkin, 1982). First, the neuroanatomy and neurophysiology of high-level vision support the proposed distinction. Two major sets of neural pathways have been identified, and the specifically visual nature of these areas has been well documented (e.g., see Maunsell and Newsome, 1987; Mishkin and Ungerleider, 1982; Van Essen, 1985). The cells in the two visual areas have different properties: The cells in area IT (in the inferior temporal lobe) tend to be highly shape sensitive (sometimes responding only to certain shapes), often are color sensitive, almost always include the fovea, and have very large receptive fields, allowing them to generalize over large regions of space (Desimone, Albright, Gross, and Bruce, 1984; Gross, Bruce, Desimone, Fleming, and Gattass, 1981; Gross, Desimone, Albright, and Schwartz, 1984). In contrast, cells in the parietal lobe are not particularly sensitive to shape or color, often do not include the fovea in their receptive fields, often are sensitive to direction of motion, and some cells in this region respond selectively to an object's location in space (as gated by eye position, see Andersen, Essick, and Siegel, 1985).

Second, behavioral data provide support for the existence of separate "what" and "where" representations and of the critical role of the temporal and parietal lobes in computing these representations. For example, Mishkin and Ungerleider (1982) tested monkeys in an apparatus that requires them to lift up a lid covering one of two food wells placed before them. The monkey's task is to learn which lid conceals food. In one version of the task, a different pattern is on each lid, each pattern being switched from the right to the left side randomly from trial to trial; the food is always under a specific pattern. In another version of the task, both lids are gray, and the problem is to learn that the relative location of a small tower (a "landmark") cues which lid conceals the food. The tower is placed closer to the lid concealing the food, with its position being varied randomly from trial to trial. If the inferior temporal lobes are removed but the parietal lobes are left intact, animals have great difficulty with the pattern learning task, but do not have great difficulty with the location task. In contrast, if the parietal lobes are removed and the temporal lobes are left intact, the reverse dissociation occurs: these animals have great difficulty with the location task but not the pattern learning task (see also Pohl, 1973; Ungerleider and Mishkin, 1982). Sagi and Julesz (1985) present convergent evidence for distinct mechanisms for "what" and "where" in humans, as revealed by psychophysical tasks that are sensitive to the distinct processing characteristics of the two systems. Gross (1978) and Holmes and Gross (1984) also showed that monkeys can learn to discriminate between objects presented at different orientations when the temporal lobes are removed but the parietal lobes are intact, suggesting that the parietal lobes encode this sort of higher-order spatial property.

Thus, the "ventral" system in the inferior temporal lobe appears to process object properties independently of location, whereas the "dorsal" system in the parietal lobe appears to process spatial properties independently of object properties.<sup>2</sup>

#### *Connections between areas*

Figure 1 displays many of the known areas involved in vision and their interconnections (from Van Essen and Maunsell, 1983), all of which provide hints as to how information is processed. (Although additional visual areas have been discovered since the time this figure was constructed [Van Essen, personal communication], it serves to illustrate the major properties of the system of importance here.) The left side of Figure 1 corresponds roughly to the ventral system, and the right side to the dorsal system. Each of these individual areas has different anatomical and physiological properties (Maunsell and Newsome, 1987; Van Essen, 1985). This illustration was constructed by observing which areas project to other areas (when appropriate stains are injected or degeneration is induced). Each area is one level higher than the level from which it receives efferent fibers, and is beneath all levels from which it receives afferent fibers (feedback). The higher the area in the diagram, the farther along it is in the processing stream.

-----  
 Insert Figure 1 About Here  
 -----

We have found the existence of several of the neuroanatomical connections particularly useful as constraints on theorizing. First, there are connections between areas MT and V4, which probably receive input primarily from the magno and parvo pathways, respectively. Both areas are retinotopically organized. Area MT has been implicated in motion perception, and cells in V4 are particularly sensitive to color and shape properties. This connection may play a critical role in using motion to derive figure/ground segregation, as will be discussed below. Second, the connections from area V4, particularly those ascending in the system, serve to define many of the regions of greatest interest for high-level vision. This distinction is particularly of interest for the highest areas. Specifically, area 7a, in the parietal lobe, and areas AIT (anterior inferior temporal) and PIT (posterior inferior temporal) will prove especially relevant in this article. Also important for present purposes, but missing from the diagram, is area STP, which is a polysensory area in the posterior superior temporal lobe, which receives connections from AIT and 7a (via the hippocampus). Third, also missing from the figure, are direct and precise connections between the regions of the parietal lobe concerned with representing location and the frontal lobe (Goldman-Rakic, 1987). Indeed, these projections terminate relatively close to Area 8, the "frontal eye fields;" this area has a role in directing eye movements. Fourth, there is a major connection, the arcuate fasciculus, between the posterior superior temporal lobe and the posterior inferior frontal lobe. As will be noted below, this



connection would allow information about object or part identity and location to be used to guide eye movements.

#### *Reciprocal connections*

To date, it has been found that every visual area that sends information to another area also receives information from that area. Thus, there are no arrow heads on the lines in Figure 1; virtually all of the connections in Figure 1 correspond to afferent and efferent pathways, with afferent pathways ascending in the diagram. Furthermore, the efferent and afferent pathways are of comparable size. Thus, a considerable amount of information flows upstream as well as downstream (Van Essen, 1985). This observation will prove useful in considering how knowledge can "prime" the visual system to search for a particular part of an object.

Although there are many other aspects of the neuroanatomy and neurophysiology that are relevant (e.g., projections from the pulvinar; the nature of chandelier cells, and so on), the ones summarized here will prove most important for present purposes. Additional neuroanatomical and neurophysiological findings will be introduced as they become relevant.

#### Processing Subsystems

Our goal is to specify what is computed by distinct *processing subsystems*, not how these subsystems actually carry out these computations (in contrast to Feldman, 1985, who attempts to grapple with both levels). A processing subsystem corresponds to a group of neurons that work together to accomplish part of an information-processing task. A processing subsystem is characterized by the input the relevant neurons receive, the operation they perform on the input, and the output they produce. The neurons that compose a processing subsystem need not be in the same anatomical location, although they probably often will be because i) nearby neurons have similar input to operate upon, ii) nearby neurons have similar outputs, and iii) nearby neurons have the opportunity for much local interaction.

#### *The hierarchical decomposition constraint*

Individual neurons often respond to more than one form of input, sometimes even crossing sensory modalities (e.g., see Bruce, Desimone, and Gross, 1981). Thus, if one's goal is to describe what neurons do, there are only two options: On the one hand, one can characterize functional systems that correspond to the individual stimulus dimensions we perceive (color, shape, intensity, texture, motion, etc.). In this case, theories of function will be cast at an abstract level, with a complex mapping from function into neurons; individual neurons would be seen as carrying out parts of numerous different functions. This is the dominant view in Cognitive Science. On the other hand, one can attempt to stay close to the brain, come what may. In this case, if an individual neuron has more than one function (using common-sense conceptions of function), such as encoding both color and orientation, the functional description will be a conjunction or interaction of some sort. In the current project, we have adopted the second approach, for the following reasons:

First, we wish to begin at a coarse level, drawing major distinctions before moving onto finer points. This seems to be the only reasonable tack, given the current level of knowledge. Although the processing subsystems we hypothesize are not primitives (the ultimate fundamental building blocks), we claim that they capture correct boundaries between distinct subsystems. If this were not true, there would be little point in working at a coarse level; the distinctions drawn here would simply be supplanted as more knowledge was gained. Thus, this approach implies an *hierarchical decomposition constraint*, which requires that further subdivision will not violate the boundaries drawn at a coarse level. Hence, we cannot hypothesize subsystems that cut across the boundaries of those we have posited; indeed, the subsystems developed here represent just such a development over those posited by Kosslyn (1987). The hierarchical decomposition constraint requires that more fine-grained subsystems either work together to accomplish the operation ascribed at the coarser level, or substitute for each other in accomplishing this operation.

Second, this requirement, like all constraints on theorizing, helps to narrow the range of possible organizations of the system. The hierarchical decomposition constraint guides us in part by constraining the appropriate level of granularity of our analyses. That is, if we are at too coarse a level (or have an incorrect theory), the hypothesized subsystems will appear to be shared in different systems. For example, language and imagery probably share numerous subsystems (e.g., ones that access stored information) and a common database. Hence, these more general terms are unlikely to describe separate neural systems, and hence will probably be of limited use in understanding how the brain processes information. The shared subsystems and database cannot be considered part of a self-contained language or a self-contained imagery system *per se*.

Third, if we obey the hierarchical decomposition constraint, we will ultimately characterize what individual parts of the brain are doing. This sort of analysis will lend the greatest insight into understanding the effects of damaging the brain. As will be discussed shortly, we would like to understand such behavioral dysfunction following brain damage in part by appeal to impaired independent processing subsystems. Requiring subsystems to decompose hierarchically is one way to ensure that the theory is describing distinct entities.

#### Subsystems of High-level Visual Object Identification

For every subsystem, we consider the information-processing that must be performed to allow a system to have the properties of the human visual system. We offer a coherent line of reasoning that led to the hypothesis we adopted. In most cases, the possible solutions were highly constrained, given the requirements of neurological plausibility and computational sufficiency. We constructed a computer simulation model for two reasons: in order to ensure that our reasoning was self-consistent and explicit, and in order to derive the implications of our reasoning (as will be described below). The purpose of this simulation was not to develop a high-performance computer vision system, but rather to consider in detail what components of high-level visual processing are necessary and what consequences follow

when the system is damaged. Following the description of our theory, we will describe the simulation, and then will turn to a review of major clinical deficits in vision that occur following brain damage; we will conclude by considering whether the theory and model lend insight into these deficits.

In order to help the reader to organize the following material, Figure 2 presents an overview of the hypothesized structure of the system at the coarsest level. Each of the major components in this figure will be decomposed into sets of subsystems as we consider what sorts of computations seem required to produce the observed behavioral abilities, keeping in mind the key constraints from the neural substrate reviewed above.

-----  
 Insert Figure 2 About Here  
 -----

#### Input to high-level visual subsystems: A visual buffer

The input to high-level visual processing is the output from low-level vision. This output is represented as a set of patterns of activation in a series of retinotopic maps (see Allman and Kass, 1976; Cowey, 1985; Van Essen, 1985; Van Essen and Maunsell, 1983). These maps preserve (roughly) the local geometry of the projection of the object onto the retina, subject to a magnification of the regions receiving projections from the fovea (see Johnston, 1986). The output from these different areas is multiscaled, with representations at different levels of resolution (cf. Campbell, 1980). This arrangement is very useful for performing a host of low-level computations (e.g., computing depth from stereo; see Johnston, 1986; Marr, 1982).

According to the criteria noted above, the boundary between low-level and high-level vision falls somewhere between V1 and V4. Moran and Desimone (1985) showed that a monkey's knowledge affects neural activity in V4, but not V1. Thus, we begin by hypothesizing that no later than V4 and no earlier than V1, one or more topographically mapped areas serves as a single functional structure, which we call the "visual buffer;" this structure receives input from low-level subsystems that detect edges on the basis of intensity change, stereo, and possibly "common fate" (area MT, which appears to process motion, provides input to V4; this input could serve to allow points moving in a common direction to be grouped as a unit in the visual buffer). We posit that this structure supports an augmented version of Marr's (1982) "2.5 D Sketch." Following Marr, we assume that local depth and orientation information are explicitly represented in this structure, and we also assume that the edge information computed in Marr's "Primal Sketch" is explicitly represented; in keeping with Marr's (1982) "principle of least commitment," it makes little sense to throw away such useful information after computing it so laboriously. The multiscale aspect of this buffer presumably reflects the existence of multiple overlapping receptive fields which differ in size and number, with the smaller, more numerous receptive fields providing better resolution (cf. Marr, 1982; Wilson and Bergen, 1979).

#### Attention window

Attention is the selective aspect of perception. Because the anatomical connections between areas in the brain are of fixed size, they can transmit only a limited amount of information in unit time. We assume that there is more information available in the visual buffer than can be passed on, and thus the system must selectively allocate the available capacity. Because of the fixed information transmission capacity, only a fixed number of retinal outputs in the visual buffer can be monitored at once. Furthermore, given that shape and location are processed separately, there must be a mechanism that splits the two kinds of information and yet keeps them tightly yoked. We posit an attention window, which will serve these functions if i) it operates by gating input from the visual buffer to the higher processing subsystems, and ii) its contents are sent to the ventral system while its location is sent to the dorsal system (cf. Treisman and Gelade, 1980).

There is good evidence that only one region of space can be attended to at a time (K. Cave and Kosslyn, in press; Downing and Pinker, 1985; LaBerge, 1983; Larsen and Bundesen, 1978; Posner, Snyder, and Davidson, 1980; Treisman and Gelade, 1980). Downing and Pinker (1985) demonstrated that attention can be adjusted in depth, which is as expected if depth is explicitly represented in the visual buffer and the attention window selects a region of the visual buffer for further processing. Of particular interest are results reported by Moran and Desimone (1985), who describe properties of IT cells that are consistent with the claim that an attention window gates input to the object properties encoding system. They found that the responses of cells in the inferior temporal lobe are greatest to stimuli at the location to which an animal is attending; the cells are inhibited when stimuli are at other locations, even when these locations are well within the cells' receptive fields. This finding can be interpreted to indicate that the "receptive field" of an IT cell indexes the range of locations of the attention window that will feed input to that cell, but at any one time the cell responds only to the current contents of the window.

If the attention window monitors only a fixed number of neurons, a simple prediction can be made: if it monitors neurons that have small receptive fields, it will sacrifice scope (visual angle subsumed) for increased resolution, and if it monitors neurons that have large receptive fields, it will gain scope at the cost of decreased resolution. And in fact, there is evidence for just such a scope/resolution tradeoff (see Egeth, 1977; Eriksen and St. James, 1986; Jonides, 1983; Shulman and Wilson, 1987). Thus, if a high-resolution encoding is required, then only a limited region of the visual buffer can be sampled. And if more than one such encoding is necessary, then serial search will be required. This claim is supported by the early findings of Sperling (1960), which demonstrated that when letters must be identified, iconic images are scanned serially in the absence of eye movements.

According to our theory, then, the contents of the attention window are treated the same way in the ventral system regardless of where the window is positioned. This property will play a role in helping one to identify objects when they are at different distances (so that their images cover different regions in the visual buffer) or are at different places in the field (so that their images fall on

different parts of the retina, and are projected to different parts of the retinotopic maps). By adjusting the attention window, similar input can be sent to the ventral system when the object's image is at different sizes and locations on the retina.

K. Cave and Kosslyn (1988) tested a simple prediction of this notion: When subjects evaluate a form, the time required should depend on the size of the region being attended prior to stimulus presentation. Thus, on 75% of the trials the sizes of two successive stimuli were the same size, but on 25% of the trials they were different sizes. And in fact, evaluation time increased linearly with the disparity between the expected and observed sizes, as expected if one has to adjust the attention window to surround the region occupied by the stimulus. This result is consistent with findings reported by Bundesen and Larsen (1975), who used a similar technique.

Only two kinds of information could possibly be used to adjust the location and scope of the attention window. First, bottom-up "preattentive" mechanisms (to use Neisser's, 1967, term) may select a region of space to be attended to solely on the basis of physical properties of the stimulus, selecting regions with distinctive color, texture, intensity, and so on. This method will be used in two circumstances: i) for an initial attention fixation in a new situation, before an hypothesis is generated about the nature of the stimulus, and ii) when an unexpected change occurs, drawing one's attention to the novel circumstance. In either case, the attention window can be adjusted without prior interpretation of the nature of the stimulus. Second, the attention window can be directed top-down, using stored information to shift it systematically in order to search effectively for a sought stimulus. This process will be described in detail below.

#### Subsystems of the dorsal system

While one is attending to a shape, the location of the shape is processed in the dorsal system. This system has at least two main stages.

##### *Spatiotopic mapping*

The "where" information in the visual buffer is retinotopic; that is, location is specified relative to the retina, not space. This representation is not useful during identification either for encoding the locations of objects in a scene or of parts of a single object, nor is it useful for navigation or tracking. Rather, one needs the location to be represented relative to objects in space, not the retina. A spatiotopic representation of the location of objects is necessary for coordinating separate objects or parts in a single frame of reference, for navigation, and so on.

Thus, there must be a subsystem that takes as input a retinotopic position, distance (computed using stereo and via other bottom-up processes), eye position, head position, and body position and uses such information to establish where an object or part thereof is located in space. These representations need not make explicit location independent of other information. Andersen, Essick and Siegel (1985) found cells in area 7a that are sensitive to the locations of objects in space as gated by eye position. This sort of conflated representation might be useful for later integration of information gleaned over

multiple eye fixations, but it would not be particularly useful for other purposes. Hence, we posit a subsystem that can be decomposed into finer subsystems that have in common the computation of location in space. These representations may well interact with each other, taking advantage of different kinds of information to converge on the best representation of location.

The dorsal system not only encodes the locations of objects in a scene, but under some circumstances will encode the locations of individual parts of a single object. That is, the principles of perceptual organization sometimes lead to parts being stored separately in memory (Biederman, 1987; Bower and Glass, 1976; Palmer, 1977; Reed and Johnsen, 1975). Encoding of individual parts is particularly likely to occur whenever one examines an object that is relatively close, so that the parts are viewed not only with high resolution but also with multiple eye fixations (with each part falling on the fovea at different points in time). In addition, parts can be encoded separately within a single eye fixation when one shifts attention covertly (Sperling, 1960). When parts are encoded separately, then, their locations must be represented in addition to their shapes.

The location of an object or part must be specified relative to something. Depending on the task at hand, different reference systems are more or less useful. For example, in order to identify a painting as being different from a subtly different fake, the locations of objects and parts should be relative to each other, the frame, or a specific point on the canvas. In contrast, in order to reach for an object, location should be relative to one's body. Thus, spatiotopic coordinates can be either viewer-centered or object-centered (Marr, 1982). (Note that retinotopic coordinates must always be viewer-centered, by definition.) Although there is some evidence suggesting that separate subsystems exist to map position using egocentric and allocentric origins (e.g., see Rizzolatti, Gentilucci, and Matelli, 1985), we have not developed this distinction in our theory.

Another issue is whether the location information should also specify the size parameters of the object. During object identification the size of a part or object is often critical; for example, one important difference between a black house cat and a panther is its size. Although we want size constancy, ignoring projected visual angle, we also want to know the actual size. Indeed, the purposes of spatiotopic mapping noted above suggest that size is intrinsically represented along with location. For example, for effectively specifying where to shift attention, one needs to know the size of the object (and in fact needs to know its distance too, so that visual angle is computed correctly prior to an eye movement). Similarly, navigation and reaching would profit if size and location are represented integrally; knowing how to avoid hitting the reckless jaywalker depends in part on knowing how big he is. Location and size are intimately related: size can be conceived of as the number of small locations an object occupies. (Indeed, local aspects of shape are no more than the distribution of locations occupied by small portions of an object, as we shall discuss shortly.) Thus, we hypothesize that the same subsystem (at a coarse level of analysis) is concerned with both kinds of information. And in fact, at

least some of the areas in the dorsal system are sensitive to changes in stimulus size (Maunsell and Newsome, 1987).

Furthermore, a parallel argument can be made that the dorsal system should encode orientation. To shift attention or navigate properly, one needs to know how an object is oriented in space. Although there is no necessary relationship between how location and orientation are actually computed, we note that logically there is a direct mapping between the two: orientation is an emergent property of representing spatial relations at different levels of scale. The orientation of a line, for example, can be represented by the relative positions of its two ends. That is, by breaking an object into parts and noting their relative locations, the orientation of the object as a whole can be computed. Indeed, the longest axis of an object in principle could be computed this way, by bisecting objects along different axes and observing which produces parts that are furthest apart. And in fact, Gross (1978) and Holmes and Gross (1984) showed that monkeys can discriminate between patterns presented at different orientations even with anterior, posterior or complete lesions of the inferior temporal lobes, thereby supporting the idea that the dorsal system is involved in encoding not just location, but also orientation.

Another issue concerns which levels of resolution should be encoded in the map. Again, it is clear that the answer to this depends on the purposes at hand. One often wants to attend to a rather precise location, as would be required to pick up a needle. One also often wants to attend to a rather coarse location, as would be useful when driving and avoiding pedestrians (one does not want to know the locations of their fingers, or even their arms in most cases). Thus, the spatiotopic mapping process must be capable of representing location at multiple levels of resolution. The multiscale representation in the visual buffer would help one to derive this representation in a relatively straightforward way.

This leads to the question of which levels of resolution should be represented at once in the spatiotopic map, which translates to the question of whether objects and parts should be represented at the same or different times in the map. On the one hand, only that which is currently being attended to (or recently attended to) might be represented. On the other hand, every parsed region in the visual buffer might be represented. On yet another hand, if you will, only those parsed regions at the level of resolution being attended to might be represented. We can eliminate the first alternative by the simple fact that one must be able to know where to look before attending to a location to encode an object. That is, without a preattentive representation that something is at a particular location, one would have difficulty in directing attention to a region likely to correspond to an hypothesized part. It is more difficult to distinguish between the second and third alternative. However, we can appeal to the same reasoning that leads us to expect the phenomenon of selective attention, namely the limited capacity of the data transmission lines. If so, then it seems likely that only regions at the level of resolution being processed by the attention window are registered in the spatiotopic map. However, this clearly is an empirical issue.

In short, we posit a subsystem (which clearly can be further decomposed) that registers the locations, sizes and orientations of all parsed units at a given level of resolution.

*Categorical relations encoding*

Following transformation to spatiotopic coordinates, one needs to encode spatial information into a long-term associative memory, where it can be used in conjunction with information being simultaneously encoded (via the ventral system) about object properties. As noted above, a fundamental property of our visual system is the ability to ignore irrelevant stimulus variations during object identification. This is particularly difficult for objects that are subject to a near-infinite number of transformations. For example, a human body can take a huge number of different postures, from fetal to standing on tiptoes with one arm held up and the other held to the side. For such mutable objects it is impossible to store a separate representation of all the possible configurations; there are too many possible positions of the parts, and new ones arise all of the time. Thus, if one simply encodes the entire object in one attention fixation, it will often fail to correspond to a stored pattern.

Clearly, it would be more useful for identification to encode aspects of objects that are invariant over their permissible transformations. Consider a human form as it contorts. Two kinds of properties do not change: no parts are added or deleted, and rather abstract spatial relations are maintained. That is, all of the limbs remain "connected to" each other in the same way, the ears remain on the "sides of" the head, and so on. Thus, it would be useful to have a representation of the spatial relations among parts that will remain constant under the transformations of part positions. An abstract, "categorical" representation specifies a class of relations, such as being "connected to," "above," "left of," or "on the side of;" members of the class necessarily have in common only one characteristic of their position, and hence such representations can capture what is stable across the various positions of such objects. Categorical spatial relations differ qualitatively from another; "above" is not a finer or different version of "inside." The categories can be relatively specific, for example by specifying the kind of "hinge" relation between the forearm and upper arm – which remains constant under all of the different positions the arm can take (cf. Hoffman and Richards, 1984).

These analyses led us to posit a subsystem that produces categorical representations of the relative locations of perceptual units (which could correspond to objects or parts). These representations are usefully combined later downstream with representations of part shapes to build up an internal model of the object. Because such representations capture general properties of a relationship without specifying the details (e.g., "next to" without specifying how much or exactly what angle), they are particularly useful for specifying the relations among adjacent parts, with each relation being relative to a specific pair of parts. This kind of "local coordinate system" is useful for building up complex descriptive structures of flexible, multipart objects (cf. Latt, Mumford, and Shah, 1984; Marr, 1982; Palmer, 1977).



The reasoning that led us to posit that orientation and size are represented in the spatiotopic mapping subsystem also leads us to hypothesize that the categorical relations encoding subsystem can be used to categorize these relations. Conceptually, there is a straightforward relationship between these different dimensions (although it is unlikely that one is actually computed from the other). If an object is broken into parts, size can be thought of as corresponding to the category of distances of the parts—larger objects will have parts that are "far" from each other, medium objects will have relations that are "medium distance" from each other, and so on. Similarly, orientation can be classified, which again can be achieved by categorizing the spatial relations of parts of the object. In this case, if one part is "above" the other, the object would be oriented roughly vertically, if one part is "to the side of" the other, the object would be oriented roughly horizontally, and so on.

One issue outstanding is whether categorical relations are computed only when one shifts attention from one part to another, or whether they can also be computed for multiple parts being attended to at the same time (i.e., encompassed by the attention window at the same time). The present theory assumes that both methods can be utilized. The assumption that relations can be computed within a single attentional fixation is grounded on the fact that relative position judgments become easier as the distance between two objects increases (obeying the Weber/Fechner law; see Schiffman, 1982), which if anything is opposite of what would be expected if serial attention shifting were required.

#### *Coordinate relations encoding*

For some objects, identification can only be accomplished by noting subtle metric spatial relations among the parts. For example, the distance between the eyes, the distance between the nose and mouth, and so on are important for identifying a specific person's face. If one views a face close up, requiring multiple eye fixations to encode it, the locations of features will be represented via the dorsal system. Simply knowing that two eyes are "next to" each other does not help one to identify a particular person; the eyes of all faces share this categorical relation. The virtue of categorical relations for identifying flexible objects is that they treat as equivalent a wide range of topographic locations, which is a drawback in cases like this. For objects that do not vary much from instance to instance and have spatial relations among parts that differ only subtly from those of similar objects, we need to represent the precise positions of parts. Thus, for some tasks a broad category of spatial relations is required, whereas for other tasks precise locations of parts are required.

Thus, we hypothesize that the dorsal system also includes a subsystem to compute "coordinate" relations among parsed regions. The spatiotopic mapping subsystem computes all locations relative to a single origin, whereas this subsystem computes relations between arbitrary pairs of objects or parts. A coordinate spatial relations representation specifies the coordinates of objects or parts relative to another object or part (allocentric coordinates) or to one's body (egocentric coordinates). These relations can in principle be specified within a "global coordinate" system, with a single origin for all objects or

parts (e.g., the body as a whole or an object in a room), or within a "local coordinate" system, with each object or part serving as an origin for another part (and hence each pairwise metric relation would be specified). Furthermore, hybrid systems in principle can be used, with objects specified relative to several small global coordinate systems, such as would occur if the locations of objects on a table were specified relative to one's hand (which would be useful for reaching) and relative to one's mouth (which would be useful for eating).

If coordinate relations are used to specify the relative locations of a pair of parsed regions, a coordinate transformation will be required if neither region serves as the origin in the spatiotopic mapping subsystem. For example, if a face is very close to one, multiple fixations will be required and the spatial relations among the parts will be computed in the dorsal system. In this case, one may want to examine second-order metric relations, such as the ratio of the distance between the eyes over their distance between the nose and mouth (cf. Diamond and Carey, 1986). In order to do so, in one encoding one eye would serve as an origin, whereas in the other the nose might serve as the origin; the coordinate relations encoding subsystem changes the origin while the same origin is used in the spatiotopic mapping subsystem.

In addition, following the reasoning offered above, we posit that this subsystem can encode quantitative measures of size and orientation. This information is necessary to encode if one is later to know where to look in the image for a specific part (as is discussed below).

Coordinate representations of spatial relations are qualitatively distinct from categorical spatial relations. That is, although "near" and "far" categories can be used in place of a coarse metric representation, these representations are not "dense;" they do not contain an indefinite number of intermediate cases, as do coordinate representations (see Goodman, 1968). Indeed, for many categorical relations there are no corresponding coordinate ones; there is no coordinate analogue to "left of," "above," "inside," and so on; these relations are independent of specific distances.

Coordinate representations are especially useful for navigation. In navigation, one often needs to know to a high degree of precision where an obstacle is located, not just that it is against a wall or next to some object. In climbing a rocky path, one wants to know how far away two rocks are, and whether the gap between them is large enough to accommodate one's foot, not just that the rocks are "next to" each other or "close together." If one examines rocks at a high level of resolution, the gap can be represented in coordinates. So too if one wants to know whether one's foot will fit into a notch in a single rock; in this case the system will parse the rock into left and right flanks, and the representations of these shapes will be processed in the ventral system while their locations, sizes and orientations are computed in the dorsal system. By increasing the level of resolution (perhaps in part by moving closer), shape can be further decomposed into subshapes and spatial relations among them. Thus, depending on the task at hand, the locations of parts of shapes can be specified more or less precisely.

Kosslyn, Koenig, Barrett, Cave, Tang and Gabrieli (in press) report a series of experiments that provides support for the distinction between the categorical and coordinate relations encoding subsystems. These experiments were motivated in part by the idea that language depends on categorical representations, and hence the categorical relations encoding subsystem might be more effective in the left cerebral hemisphere (along with most other language-related processing; see Hecaen and Albert, 1978). In contrast, navigation depends on coordinate relations, and hence coordinate relations encoding might be more effective in the right hemisphere (at least one component of which appears to be more effective in the right cerebral hemisphere; e.g., De Renzi, 1982). (The actual motivation for this prediction is more subtle, but these general ideas serve to provide the gist; see Kosslyn, 1987, for details of the motivation). Kosslyn et al. found that left-visual-field/right-hemisphere presentation of simple tasks involving metric relations (such as deciding whether an X and an O are greater than or less than one inch apart) resulted in faster and more accurate performance than right-visual-field/left-hemisphere presentation. And vice versa for simple tasks requiring categorical relations (such as deciding whether an X is left of an O). A left-hemisphere superiority was found for three categorical relations (left/right, above/below, on/off) and a right-hemisphere superiority was found when three different distances were evaluated (ranging from 2 mm to 2.54 cm). This dissociation, then, provides evidence for the distinction between the two subsystems, above and beyond the evidence of hemisphere differences.

It is of further interest that Kosslyn et al. found that after much practice at a metric judgment task, the right-hemisphere advantage disappeared. However, when Koenig, Gabrieli, Kosslyn, and Lin (1988) repeated the experiment, replicating the result, they also brought subjects back and tested them again on the following day. They found that the right-hemisphere superiority for the metric task was reinstated at the beginning of the trials. (Apparently, categorical spatial relations representations are not consolidated overnight; changes in neural connections during initial learning apparently are quickly lost, and repeated use is necessary for permanent alteration within the neural network that comprises this subsystem.) For present purposes, the important implication of this result is that categorical relations representations are not simple recodings of metric judgments. A distinct neural subsystem seems to underlie this processing.

These results are consistent with many other findings in the literature. For example, Taylor and Warrington (1973), Warrington and Rabin (1970), and Hannay, Varney and Benton (1976) all found that right-hemisphere damage disrupts dot localization more than left-hemisphere damage (but also see Ratcliff and Davies-Jones, 1972, for a failure to replicate using an easier task). Similarly, Hock, Kronseder and Sissons (1981) found that only the right hemisphere shows orientation dependence when figures are judged as being the same or different. This result would follow if the left hemisphere assigns categorical relations that are invariant over orientation (e.g., "connected to;" see also Mehta, Newcombe, and Damasio, 1987; Olson and Bialystok, 1983).

In summary, the dorsal system includes an early subsystem that computes locations of perceptual units (corresponding to objects or parts) in spatiotopic coordinates, and then two more subsystems that operate in parallel on this output; one of these subsystems computes categorical relations between perceptual units whereas the other computes coordinate relations between units.

#### Subsystems of the ventral system

We argue that the ventral system can be decomposed into three types of subsystems. The rationale for this decomposition is presented below.

##### *Preprocessing*

As noted earlier, among the basic properties of our visual systems is the ability to identify objects when their images subtend different visual angles, when they are seen from novel vantage points, and when they fall in different places in the visual field. Because the size of the attention window can be adjusted, it can envelop shapes that occupy different areas in the visual buffer (i.e. that subtend different visual angles), and because the location of the attention window can be shifted, it can encode patterns in different locations within the buffer. However, although these properties help the system to identify objects when they appear at different sizes and in different parts of the field, they do not in and of themselves confer the ability to identify objects in these varying circumstances. Indeed, we can identify objects when they subtend angles so large as to require multiple eye movements, and hence the attention window can never envelop the entire object. Furthermore, we can identify objects when we see them from novel vantage points.

Lowe (1987a, b) points out that certain aspects of an object's image remain relatively constant under scale changes, rotation, and translation. Lowe noticed that although these aspects of the image are not precisely the same in different conditions, they are similar enough from case to case to be unlikely to have arisen from chance. For example, parallel edges of an object tend to project roughly parallel lines, no matter how the object is aligned (although precise parallelism is disrupted with perspective, but not very much if the object is relatively small). Similarly, places where edges intersect will project intersecting lines (except in the degenerate case where they are superimposed), parts that are close together will tend to project edges that are close in the image, symmetrical parts will tend to project symmetrical patterns, and so on. Lowe (1987a) proposes a Bayesian method for estimating the probability that a given instance of these properties is due to chance. Lowe argues, as does Biederman (1987), that these "nonaccidental" image properties are extracted and used to access stored representations of shape (Biederman, 1987, presents a good summary of Lowe's nonaccidental properties).

The preprocessing subsystem we posit uses a combination of edge, texture, color, and intensity information to locate the nonaccidental properties on the input (the contents of the attention window). We call these nonaccidental properties "trigger features." The trigger features are only weakly

revealed by any one source of information, and hence much computational mileage is gained from using combinations of different sorts of information.

The trigger features are, by definition, rather impoverished relative to the image itself; they are features of the image that are likely to remain constant under different viewing conditions. In many situations, the trigger features will not be sufficient to implicate one object uniquely, particularly when parts of objects are occluded; as Lowe (1987a) notes, in these cases input must be compared to a stored image. Furthermore, the image itself must be encoded downstream, if only because mental images require that these patterns be stored. Thus, we posit that the preprocessing subsystem marks the trigger features on the image itself, thereby encoding both the trigger features and the image into the system. (In our computer simulation, to be discussed below, we literally place asterisks along parts of the edges that correspond to trigger features, and later examine the pattern of these asterisks.) This marking process is entirely driven by the "nonaccidental" properties of the stimulus; it is done bottom-up, prior to identification.

The patient described by Riddoch and Humphreys (1987) and Humphreys and Riddoch (1987) appears to have a deficit in this subsystem, being able to represent only a few trigger features at one time. Thus, when looking at an entire object, the trigger features were not sufficient to allow him to identify the object. This patient had selective difficulty in identifying overlapping figures, naming line drawings and objects, and in determining whether a shape corresponds to an object. However, he was able to judge whether a silhouette shape corresponds to an object better than he was able to judge whether a line drawing corresponds to an object, which makes sense if the line drawings added additional trigger features that taxed the system (but which were not distinctive for a given object). Furthermore, he was able to determine whether two drawings depicted the same object (even when seen from some different points of view), could draw reasonably well, and could identify many individual features of objects (e.g., an elephant's legs). As Riddoch and Humphreys point out, these tasks can be done piecemeal, with encodings of local parts of the shape being compared sequentially. In this case, the trigger features would be allocated over a smaller region, and fewer such features would be necessary to mark; thus, if the preprocessing system were limited in the number of such features it could mark at the same time, it would have an easier time encoding parts one at a time.<sup>4</sup>

#### *Pattern activation*

Visual object identification requires that input be matched against previously stored information. The system must include memory representations of previously seen objects. The pattern activation subsystem we hypothesize contains modality-specific representations that specify visual properties of previously seen shapes. These representations are matched against input, and recognition occurs when a sufficiently close match occurs. When recognition occurs, an output is produced that serves to convey a classification representation to later stages of processing. In order to ignore irrelevant shape variations, this subsystem must be capable of producing one output from a range of similar inputs.

Given that such generalization is desirable, however, one is faced with the problem of how to define the range of acceptable generalization over variation in the input. The most straightforward approach is to have the output indicate not only the stored representation that best matches the input, but the degree to which it uniquely implicates that object or part. Thus, if the match is poor or more than one representation matches the input to a similar degree, higher levels of processing will not take the object to have been firmly recognized, and will engage in further processing (as described below). The output from the pattern activation subsystem, then, can be regarded as a kind of name (but which is not in a natural language) with a confidence rating.

The pattern activation subsystem we hypothesize is modality-specific. Thus, it can play only a limited role in object identification, given that knowledge about objects can be addressed via multiple sensory modalities and is often amodal (e.g., the name of its category, abstract properties of the object such as its value, and so on). The pattern activation subsystem we hypothesize only associates visual properties with a classification representation. As will be discussed shortly, in some circumstances the output from this subsystem is sufficient for identification to occur downstream, but under many circumstances it is not. In either event, complete identification occurs only when the entire range of stored information associated with the object is accessible, which (if only because some of this information is not modality-specific) must occur further downstream from the hypothesized pattern activation subsystem.

Miyashita and Chang (1988) present evidence that cells in the anterior inferior temporal lobe represent visual memories. They found cells in this area that responded selectively to particular stimuli during a retention interval, after the stimulus had been removed. This result is consistent, of course, with the data indicating that visual memories are disrupted when the temporal lobes are removed (e.g., Ungerleider and Mishkin, 1982).

Following Lowe (1987b), we note that the "trigger features" (which we posit are computed in the preprocessing subsystem) and their relative positions often are mutually consistent with only a single shape as seen from a single point of view (Lowe calls this the "viewpoint consistency constraint"). Thus, the trigger features alone may be sufficient to identify an object when it is seen at different sizes and positions in the visual field. However, when an object is partially occluded, or in unfamiliar orientations in depth, these cues alone may not be uniquely consistent with a single shape; more than one shape may be consistent with the trigger features and their positions. (Biederman, 1987, suggests that shape representations are composed of geometric primitives, each of which is accessed by a set of trigger features, but we need not specify these aspects of the representations here.)

When the trigger features do not strongly implicate a single object, it is useful to activate a stored representation of the shape *per se*. The image itself (as encoded via the attention window) can then be compared to the pattern stored in memory (in the pattern activation subsystem, according to the present theory). The additional information in the pattern may be adequate for distinguishing among

alternative objects or parts that are consistent with the trigger features. The sizes, locations, and orientations of these stored and input representations can be adjusted until a best match is found (cf. Lowe, 1987a; Ullman, 1986). This procedure allows one to have the best of both worlds—the robustness of the trigger features and the richness and detail of the projected shape.<sup>5</sup>

In order for this scheme to work, the long-term memory representations must be organized in such a way that viewer-centered projections are accessed for comparison to input. However, the present theory leaves open the question of whether the stored representations themselves are viewer-centered (with different representations for different points of view), or object-centered (with different projections being accessed). Perrett, Smith, Potter, Mistlin, Head, Milner, and Jeeves (1985) describe cells in the superior temporal sulcus (STS, which defines the upper boundary of IT) whose behavior bears on this issue. About 10 - 11% of these cells respond selectively to static views of the head, and Perrett et al. found some of these cells are tuned very narrowly, responding to the head in specific orientations or to the eyes in specific positions (different directions of gaze). In contrast, other cells were found to be tuned more broadly, responding to the head or eyes across a number of different orientations or positions. Some 69% of the cells that were selectively responsive to one class of object (the face or head) responded best when the object was seen from a particular vantage point; in an extreme case, cells responded preferentially to a single profile (i.e., a face seen from either the left or right side). About one-quarter of the cells that responded preferentially to faces were relatively insensitive to viewpoint. These data indicate that the input is often compared to viewer-centered representations; however, they do not tell us whether such representations are stored or are projections from a richer underlying representation. Nevertheless, it seems clear that we must reject Marr's (1982) view that fully three-dimensional, object-centered representations are always compared during object recognition and identification (for further discussion of this point, see Perrett et al., 1985; see also Jolicoeur and Kosslyn, 1983).

Although the mechanism proposed by Lowe is very useful for identifying rigid objects, it will not lead to identification of flexible objects in unusual configurations (e.g., a contorted person or tumbled bicycle). However, the same mechanism can be used to identify individual parts (which also correspond to patterns stored in the pattern activation subsystem), which may be rigid although the object as a whole is not. Indeed, the shapes of many individual parts tend to vary relatively little from instance to instance. For example, although an image of a person can take numerous configurations, the shapes of many of the individual segments (forearms, fingers, heads) do not vary so widely. A shape can be parsed into separate parts when it is viewed close up, with high enough resolution so that parts are clearly visible (cf. Bower and Glass, 1976; Palmer, 1977; Reed and Johnsen, 1975). Depending on the level of resolution selected (by top-down mechanisms, as described below), the same stimulus often can be encoded as a single, lower-resolution whole or as multiple higher-resolution parts. Indeed, Perrett et al. (1985) found cells in STS that responded preferentially to the eyes per se; these cells responded as

well to the eyes viewed through a slit as to the entire face. Other cells, in contrast, did not respond to the eyes alone, but did respond well when the entire head was viewed.

Thus, a subsystem that matches input to stored shapes can contribute to allowing us to ignore irrelevant shape variations if we make two assumptions: First, shapes that do not match a stored representation well are decomposed into their constituent parts, which are then processed separately in the ventral system. Second, separate representations are stored for each distinct type of part (e.g., there may be five different prototypes for the shape of the back of a chair). The cues extracted by the preprocessing subsystem can be used to access representations of parts as well as wholes (cf. Ullman, 1987). If so, then it will be relatively easy to generalize over different examples of parts, because the trigger features from the preprocessing subsystem (which themselves strip away irrelevant variation) will access the appropriate stored representation in the pattern activation subsystem.<sup>6</sup>

It may be tempting to take Perrett et al.'s (1985) finding that some cells are sensitive to direction of gaze and head position, whereas others are not, as evidence that coordinate and categorical relations are used in the pattern activation subsystem. We can reject this idea for a number of reasons: First, if such representations are present, why would ablation of the parietal lobes severely impair an animal's ability to learn spatial discriminations and to make spatial judgments? Second, if such representations are used in the pattern activation subsystem, either they would be computed redundantly along with those in the dorsal system, or there would be direct projections from the parietal lobes to IT. No such pathways are presently known. Third, there is a much simpler interpretation of Perrett et al.'s results, which rests on the observation that the parallel to the distinction between categorical and coordinate relations encoding lies in differences in the precision of shape categories. That is, these results indicate that some cells have sharper gradients than others. (Such generalization gradients are a natural byproduct of computation in neural networks [Rumelhart and McClelland, 1986], and—as noted earlier—we assume that each of our subsystems corresponds to a neural net.) Thus, we do not posit separate subsystems for different ranges of input because i) there is presumably a continuum of bandwidths, and ii) because the differences are quantitative, not qualitative, the subsystems would presumably receive the same input, perform the same qualitative type of operation, and produce the same type of output. Hence, by our definition of what characterizes a subsystem, the subsystems would be the same.

C. Cave and Kosslyn (1988) tested a key prediction of this theory in the following way: Line drawings of common objects were disrupted either by cutting objects up into parts at the natural parse boundaries (as determined by subject ratings) or by cutting them up at arbitrary locations, violating natural parse boundaries. The fragments were then either exploded outwards, maintaining their relative spatial positions, or were scrambled, disrupting the spatial relations. (An additional group of subjects rated each picture for degree of overall "disruptedness," and this factor was controlled.) The present prediction, based on the importance of the viewpoint consistency constraint posited by Lowe



(1987b), is that the disruption of the parts should not be as important as the disruption of relations: When relations are preserved, the trigger features will still be in the correct relative positions. Although they will define a stretched or slightly distorted object, the pattern activation subsystem should be capable of generalizing over such relatively minor distortions. However, when the relations are disrupted, the trigger features are in incorrect locations. And in fact, this prediction was borne out: Disrupting spatial relations resulted in severe impairment of naming response times (as measured by having subjects speak responses into a voice-activated relay). Indeed, there was very little effect of disrupting part boundaries when relations were preserved.

#### *Feature detection*

Not all of vision is dedicated to representing shape. Our judgments of aesthetics, for example, depend on encoding color, texture, and other aspects of stimuli that are not simply matched to attributes stored in memory. Although we posit that the preprocessing subsystem uses color, texture, and intensity information, we claim these dimensions are combined to help discover nonaccidental properties of the shape. We also claim that at least color (and probably other dimensions as well) is processed by subsystems that send this information directly to associative memory, not to the visual memory per se. (It is tempting to suspect that this subsystem receives information from the blob areas in V2, which are part of the parvo system.) Such properties presumably allow one to make judgments about properties of objects (e.g., when judging whether a melon is ripe), and could be used to decide whether two visible shapes are the same or different. Thus, we posit a very coarse subsystem, which clearly can be decomposed further, that extracts features (for lack of a better term).

#### Associative memory

Object identification requires accessing stored information associated with an object. This information typically includes facts about the object's name, categories to which it belongs, familiar contexts in which it is found, names of other objects that are frequently encountered with it, its functions, its cost, its constituent parts and their spatial relations, and so on. The stored information needs to be accessed via multiple sensory modalities; for example, one can identify a cat by seeing it, hearing it, or feeling it. Furthermore, the stored information is often amodal. Thus, we need a memory representation that is not modality specific, but that associates modality-specific information with other sorts of stored information. This representation would serve to associate the relevant information with an object, and hence (directly or indirectly), with each other.

Thus, we posit that the outputs from the object properties and spatial properties encoding subsystems are passed to an associative long-term memory, where the two types of representations are conjoined. We argue that there must be such an integrated representation for a variety of reasons. For one, humans can recall where objects belong in a scene and where individual parts belong on an object, and hence there must be a locus in which shapes are associated with locations. For another, as

Attneave (1974) pointed out so persuasively, the arrangement of parts is an important aspect of shape, and hence we expect information about shape and location to come together at some stage in processing.

This structure probably corresponds (at least in part) to an area in the temporal lobes, possibly near the human analog to area STP (short for superior temporal polysensory) in the posterior superior temporal lobe (e.g., see Bruce, Desimone and Gross, 1981). Area STP would provide some initial processing that would be useful for an associative memory. Over half of the cells in this area respond to input in more than one modality; cells in STP receive converging input from visual, auditory, and somesthetic systems (from IT, superior temporal auditory cortex, and from posterior parietal cortex). STP is not topographically organized. Bruce et al. found that some of these cells (45% of those that responded to visual stimuli) were selective for particular stimuli (e.g., faces), that cells in this area had very large receptive fields (most being over 150° of visual angle), that responses were equivalent across the receptive field (unlike IT neurons, which typically respond better to foveal input), that many cells were directionally sensitive and that most responded best to moving stimuli. Some cells responded best to a complex combination of visual and auditory input. The cells were not sensitive to size, orientation, or color. (Perrett et al. (1985) note that some of the face-specific cells they studied may have been located in STP; if so, it would be of interest to know whether these particular cells were tuned for faces seen from particular viewpoints.)

All of these properties suggest that STP is receiving input that has already undergone modality-specific processing. Indeed, the anatomical connections to IT and the parietal lobe, in conjunction with the responsiveness of these cells to both shape and movement (and possibly also location, which was not tested), is consistent with our notion that the ventral and dorsal inputs converge here. Roughly speaking, the human analog to this area would appear to be near Wernicke's area (in the posterior, superior temporal lobe), which appears to be involved in representing information used in language comprehension (e.g., Hecaen and Albert, 1978).

When the overall shape has been closely matched in the pattern activation subsystem, the classification and confidence level output from that subsystem (which serves as input to associative memory) may be enough to implicate a single stored data structure in associative memory, resulting in object identification. In this case, the spatial properties (where the object is located relative to the viewer or another object) would not contribute to object identification. However, if an object is viewed over the course of multiple eye fixations, with separate portions being encoded during each, object property and spatial property inputs must be integrated in associative memory to identify the object.

Thus, one goal of processing in associative memory is to use object property and spatial property inputs to access appropriate stored information, leading to object identification. We can conceive of this as a constraint-satisfaction process; the system tries to converge on what object is being seen by finding the stored representation that is most consistent with the object property and spatial property information being encoded. We hypothesize that objects are represented in associative memory by

amodal, "propositional" structural descriptions (because they are amodal they can be addressed by sensory input from multiple modalities). These representations indicate the parts and characteristics (such as color and texture) and their spatial relations (Latto et al., 1984; Marr, 1982; Palmer, 1977). During identification, we hypothesize that encoded object properties (parts and characteristics) and spatial relations (from the ventral and dorsal systems, respectively) are matched in parallel to properties and relations of stored representations. To the extent that the input properties and their spatial relations match those of a stored object, one can be confident that one is seeing that object; and to the extent that properties and spatial relations are distinctive for one object, one can reject hypotheses favoring other objects.

The ability to generalize to new shapes of an object cannot be a consequence solely of breaking an object down into relatively invariant parts and categorical relations between them. Some members of classes differ in the presence or absence of parts or properties (e.g., arms for chairs or spots for dogs). One way of coping with this problem is to set a threshold in associative memory. That is, if each property and corresponding relation is viewed as evidence, we can assign a weight to each one. An object is identified when enough weights have accumulated to exceed the threshold – regardless of which properties and relations contributed to the weights. Thus, many different combinations of properties and weights will allow the object to be identified. For example, if a "chair" is characterized by a seat (very important, very high weight), legs (important, medium weight), a back (important, medium weight), arms (not very important, low weight), and so on, we can identify a wide variety of chairs if the threshold is set so that all we need are a seat and two or more of any of the other properties (see Smith and Medin, 1981; Wittgenstein, 1953). When properties and relations are encoded that are inconsistent with an object (e.g., a lightbulb, suggesting that one is viewing an unusual lamp, not a chair), its threshold is raised (and another possible hypothesis that is consistent with the new input is formulated and tested, if the system is engaged in top-down hypothesis testing, as described below).

Finally, depending on the task at hand, different sorts of information will be relevant in associative memory. This observation is definitional, and has some important implications. For example, if one is shown a face and asked to name a particular person, many possible responses ("face," in this case) must be inhibited. One does not want to know merely that the object is a face; one wants to know which face it is. Thus, the system must set itself to allow only representations with specific indexing features (e.g., level of specificity) to remain activated. Furthermore, this function would be even better served if there was feedback to the pattern activation subsystem, priming appropriate representations and inhibiting inappropriate ones. Indeed, such a feedback mechanism is needed to activate stored patterns in mental imagery, because i) we can form mental images upon being given the name of the to-be-imaged object, and ii) the pattern activation subsystem is the only location where

visual memories are stored. Such feedback could be one purpose of some of the descending pathways in the visual system.

#### Subsystems used in top-down hypothesis testing

The eye encodes only relatively little high-resolution information at any given fixation (about 2° of visual angle). Thus, identifying many objects or scenes will require multiple fixations, and attention must be shifted to new locations. Two kinds of information can be used to direct attention. First, changes in stimulus input – such as a flashing light or sudden movement – can draw one's attention to a specific location. This "bottom-up," stimulus-driven mechanism is apparently responsible for attention shifting in young infants (e.g., see Bower, 1970). Second, knowledge, belief, or expectation can be used "top-down" to drive sequences of attention shifts. Yarbus (1967) provides ample evidence that eye movement patterns often reflect the use of knowledge about the inspected picture. Particular cells in the parietal lobes seem to be involved in this process; these cells show increased activity immediately before a voluntary attention shift, and do not show increased activity if attention is not voluntarily initiated by the animal (e.g., see Lynch, Mountcastle, Talbot, and Yin, 1977; Yin and Mountcastle, 1977). In addition, Colby and Miller (1986) found that some cells in STP are time-locked to the initiation of a saccade, which is consistent with the notion that these cells have a role in controlling where the eyes will move.

Some shifts of attention, then, are based on stored knowledge, which can be used to formulate and test hypotheses about what we are seeing. Not all hypotheses need be specific, such as that one is viewing a cat and hence one should be able to find whiskers at the front of its face. If the input does not implicate a single object, one can look to one side of the object for a distinctive property. This sort of weak hypothesis testing is a default strategy that is based on a weak heuristic that important properties are on the front or back of objects, and is better than totally random search. Any sort of directed information-gathering is more efficient than unsystematic encoding and waiting until enough information is encoded to discover what it is we are confronting. Thus, we hypothesize that as property and location information enter associative memory, we actively generate an hypothesis (or hypotheses) of what the object is and then look for properties that should be present if we are correct (cf. Gregory, 1970; Neisser, 1967, 1976). As properties and relations are encoded, some object representations will be better satisfied by the input, which presumably leads them to become more highly activated. As an object representation becomes more highly activated, the properties associated with the object (which are integral components of the object representation) in turn become activated. In our use of the term, the greater the "activation," the more easily information-retrieval processes (to be discussed shortly) can access the representation. It is important, then, that we posit that the more often one uses a particular fact, the "stronger" the representation of that fact becomes, and that stronger representations require less activation for the information to become accessible. By definition, distinctive properties are those that serve to distinguish an object from other similar objects.

Thus, these properties should be used disproportionately often, and hence should become stronger and more easily accessed.

We hypothesize that attention is directed to the locations of the properties of the most activated object, in an effort to discover whether those properties do indeed characterize the stimulus. We will focus here on the testing of specific hypotheses, which is the more general case (the default strategy is the same except that it does not vary for different inputs, and so is less interesting.) As will be developed below, attention is directed by moving the attention window to the correct location (which sometimes may require prior moving of the head or eyes), and priming the appropriate representation in the pattern activation subsystem, biasing it to categorize the input that way.

Unlike the subsystems described above, the following subsystems do not constitute components of a single subsystem characterized at a coarser level; the subsystems described below are not used only in the service of carrying out top-down search, and hence this component violates the hierarchical decomposition constraint. Thus, we would not expect to find neural tissue corresponding to a distinct "top-down hypothesis testing" subsystem in the brain, but would seek the individual component subsystems themselves. We have grouped these subsystems together here merely for expository purposes. Figure 3 illustrates all of the processing subsystems posited by the theory, and indicates the flow of information used in the subsystems described here as well as those discussed previously.

-----  
 Insert Figure 3 About Here  
 -----

#### *Coordinate property lookup*

In order to test an hypothesis, we first must be able to look up in associative memory the properties an object should have. Such a subsystem would be most efficient if it began by looking up particularly distinctive properties. That is, if one wanted to tell a cat from a dog, looking for four legs would not help much. But looking for a particularly shaped head would be helpful, as would looking for vertical slits in the eyes. In order to use such information to test an hypothesis one must know where to look for the property. We argued above that location is represented in two ways, using categorical or coordinate relations. Thus, we posit another subsystem that accesses properties for which location is stored as coordinates. This sort of representation is useful for identifying rigid objects. For example, the mouth of the Mona Lisa is always in exactly the same place within the frame, and hence one can usefully store a coordinate representation that directs attention to the appropriate location. Because spatiotopic coordinates are used, which can be object-centered, this sort of representation will be useful even when rigid objects are seen from different points of view.

#### *Categorical property lookup*

We also hypothesize a subsystem that looks up properties specified using categorical relations. The reason we posit two distinct lookup subsystems is that there are very different computations to be

performed if coordinates or categorical relations are accessed. If a coordinate relation is looked up, we are in luck – we simply look in that location. But if a categorical relation is looked up, we must somehow convert this information to a region in space, a range of coordinates.

In addition, although coordinates can be specified relative to a single origin (e.g., the center of a picture), categorical relations are necessarily relative to another property or object. Thus, in order to know where to look, one first must find the reference property. That is, categorical relations incorporate local coordinate systems, with each property being specified relative to another. For example, a cat's "head" might be specified as being "connected to the top of the neck," and its neck might be specified as "connected to the front of the body." If so, then we first must locate the part serving as a reference point. For example, if we are currently focused on the feet, a chain of such connections must be made (the feet are connected to the ankle, which is part of the foreleg, which is connected to the thigh, which is connected to the body). In our simulation model, this process consumes a surprising amount of computation, none of which is necessary if an appropriate coordinate representation is used.

We assume that both lookup subsystems operate at the same time, and that they are mutually inhibitory. Thus, if both succeed in finding relevant information, the one that accesses the "stronger" information will inhibit the other. If the coordinate property lookup subsystem wins, it sends the coordinates to the attention shifting subsystem while at the same time it sends a visual code (naming the property itself) to the pattern activation subsystem in order to prime it for the expected part. If the categorical property lookup subsystem wins, it sends the categorical relation to the categorical-coordinate conversion subsystem while sending a visual code (naming the property itself) to the pattern activation subsystem in order to prime it for the expected part.

#### *Categorical-coordinate conversion*

Whenever one looks up a categorical relation, the category must be converted to a specification of a location in space. Because this is a very different task than looking up the representation, we posit a distinct subsystem to carry it out. The actual computation of the coordinates corresponding to a categorical relation is surprisingly (to us) complex. One needs coordinate information about the size of the object, its taper (or equivalent information to be used to determine front and back), and its orientation. Even given this, one will not be able to specify the location precisely. Thus, we have posited that this subsystem initially computes a range of coordinates, and then uses information about the locations of perceptual units in the image (encoded by the coordinate relations encoding subsystem, via associative memory) to direct attention to the proper location.

That is, we hypothesize that this subsystem produces an initial set of approximate coordinates by an "open loop," unsupervised heuristic procedure. Such a process is relatively fast and requires less effort than carefully guided (as opposed to ballistic) movement. But because it is very difficult to ensure that such a process actually "zeros in" on the proper location, once attention is near a perceptual

unit, a "closed loop" process is used. In this second phase, the categorical-coordinate conversion subsystem directs the attention shifting subsystem to the coordinates of the object or part nearest to the attention window. Limiting this closed-loop process to a "fine tuning" role minimizes its additional expense, both in time and effort, while still taking advantage of its capability for high precision.

#### *Attention shifting*

There must be a subsystem that actually shifts attention, adjusting the attention window, the eye, head and body, as appropriate. Furthermore, the attention shifting subsystem must be capable of altering focus in depth; we assume that depth information is implicit in the visual buffer, and hence regions corresponding to input at different depths can be selected as well as different regions in the plane.

The property lookup subsystems must send spatiotopic coordinates as instructions to this subsystem, given that those are the only kinds of coordinates that are stored (recall that location is encoded via the categorical and coordinate encoding subsystems, which in turn operate on output from the spatiotopic mapping subsystem). However, the attention shifting subsystem must be capable of shifting the attention window in the visual buffer, which requires specifying retinotopic coordinates. Thus, the attention shifting subsystem must be capable of a coordinate transformation, computing the inverse of the mapping function used in the spatiotopic mapping subsystem. The second phase of instructing the attention shifting subsystem, in which feedback guides fine-tuning the location of the attention window, is useful in part because it simplifies the complex coordinate transformation needed to shift the attention window in retinotopic coordinates on the basis of spatiotopic instructions; the "shift to the nearest parsed unit" strategy does not require specifying a target location in high precision in advance.

The attention shifting subsystem can be decomposed into at least three more fine-grained subsystems. Posner, Inhoff, Friedrich, and Cohen (1987) hypothesize a subsystem that shifts attention to a position in space, another subsystem that engages attention at that position, and a third subsystem that disengages attention when appropriate. The subsystem that shifts attention appears to involve the superior colliculus, the one that engages attention appears to involve the thalamus, and the one that disengages attention appears to involve the parietal lobes. We have implemented this mechanism only very coarsely in our simulation models, specifying only a single attention shifting subsystem.

#### Summary

In order for an object to be identified, its image must be projected into the visual buffer. The attention window gates which information in the visual buffer is sent to the dorsal and ventral systems for further processing. Object properties are processed in the ventral system. The preprocessing subsystem uses shape, color and texture to extract nonaccidental "trigger features" and marks them on the image. These features are then matched against those of previously encoded objects in the pattern

activation subsystem. If these trigger features match poorly, the entire image is matched to stored patterns. Other features of the object are encoded directly into associative memory by the feature detection subsystem. At the same time this processing is occurring, the location, size, and orientation of the object are sent to the dorsal system, which represents the location relative to some other object or part. The spatiotopic mapping subsystem computes the location in space, actual size, and orientation, and the categorical and coordinate encoding subsystems encode these types of spatial information into associative memory.

If the match to a stored shape is very good in the pattern activation subsystem, the output from the ventral system will be sufficient for a unique identification in associative memory. However, if the match is not optimal, then only a tentative match is made, leading to an hypothesis to be tested. This is likely to occur if the object is viewed under impoverished conditions (including being so close that not all of it can be encoded in one fixation) or if an unusual version of the object is viewed. During hypothesis testing, stored properties (parts and characteristics) of the candidate object and their locations are accessed by one of the property lookup subsystems. The location of the most strongly activated property is used to shift the attention window to a new position and, typically, level of resolution. The shape and other object properties of the portion of the object found at that position are sent to the ventral system, and the location and other spatial properties of this portion of the object are sent to the dorsal system. The new dorsal and ventral inputs are processed by the relevant subsystems and yield a new input to associative memory. If this input to associative memory is consistent with the properties of the candidate object, then one has evidence in favor of that hypothesis. The amount of evidence that is necessary for identification depends in part on how distinctive the object's properties are and in part on the context. If partially confirming evidence is not found, another possible hypothesis that is consistent with the new input is formulated and tested. This top-down hypothesis testing cycle is repeated as many times as necessary until an object has been confirmed.

The set of subsystems and interconnections we have hypothesized are illustrated in Figure 3, and a summary of the properties of the individual subsystems is presented in Table 1.<sup>7</sup>

-----  
Insert Table About 1 Here  
-----

## II. DERIVING PREDICTIONS USING A COMPUTER SIMULATION

Our aim is to understand how the processing subsystems operate in concert during perception, both in the normal and damaged brain. According to our characterizations, each subsystem is dependent on input from specific subsystems and produces output to yet other subsystems. Thus, the interactions among the subsystems are quite constrained. In this section we begin by outlining how the subsystems are used to perform four different types of tasks, with several special cases of each type. We then turn to the ways in which damage affects the operation of the system. Following this, we consider the ways



in which behavioral dysfunction arises from damage, and then we describe the results of selectively damaging our model of the system and observing it perform the tasks. The output from the program makes several interesting, in some cases counterintuitive, predictions. In the final major section we will review the major types of visual deficits that actually occur following brain damage, and consider whether our theory and model provide insight into these maladies.

### The Computer Simulation Model

The theory of information processing summarized above is complex, and thus it made sense to implement the theory in a computer simulation model. Indeed, without such a model it is very easy to fudge implausible predictions by hand-waving, and to formulate likely predictions by creative interpretation of the theory. Given the existence of a running computer simulation model we are assured that: i) the theory is not vague; ii) the theory is not internally inconsistent; iii) the theory indicates clear directions in which one can account for many of the basic functions of the intact system; and iv) specific predictions can be derived. We will provide a brief overview here, which should be sufficient to understand how the predictions were generated.

Our computer simulation model is not a computer imaging system that performs both low and high-level vision functions on digitized camera input; instead the program bypasses low-level vision to compute high-level vision functions on input consisting of hand-segmented numerical arrays. Furthermore, we are interested in the ways subsystems interact, not in how they actually process input. Thus, in many cases we have taken shortcuts to make the subsystems operate in as brief a time as possible, sacrificing generality while doing so. (For example, the program only identifies two-dimensional pictures, and it has no representation of the third dimension. Furthermore, trigger features are marked on an *ad hoc* basis, and we presegment the input images into parts so that the program can examine trigger features on a single part—the one that fills the major portion of the attention window after attention has been shifted to the location of a sought part—when encoding a part; we have not tried to duplicate Lowe's work.) The purpose of the simulations is to derive predictions of the theory, not to process actual images. Thus, we developed a system that was capable of performing highly simplified versions of the tasks, which were sufficient to allow us to observe the effects of damage.

The individual subsystems were implemented as separate functions or groups of functions. The input to the system consists of 60 by 60 arrays that are meant to represent pictures that have been parsed by lower vision processes, different parts being represented by different numbers.<sup>8</sup> These arrays are placed directly into the visual buffer. According to the theory, connections between subsystems are capable of transmitting less information than is contained in the visual buffer. We simulate this by using a 20 x 20 array to represent the information reaching the dorsal and ventral systems from the contents of the attention window, with a data line corresponding to each cell. Thus, the capacity of the data line from the visual buffer is 400 characters. This capacity limitation results in a tradeoff between scope and resolution: depending on the size of the attention window surrounding the area of

interest, the 20 by 20 array can represent the whole image at such low resolution that part segmentation is not discernible, or it can represent segmented image sections in high or medium resolution. Patterns in a 20 x 20 array are used as the input to the dorsal and ventral systems. The dorsal system uses the parsed regions in the input array and the size (relative to the visual buffer) and location of the attention window to compute location information, and the ventral system uses the input array to compute pattern-match information. The "where" information from the dorsal system and the "what" information from the ventral system are transmitted to associative memory where they are placed on an object short-term memory structure. Resident in associative memory are long-term data structures containing information about the objects. All associative memory structures are implemented as property lists.

In the version of the simulation used to generate the predictions described below, only eight stimulus pictures were used; although the theory posits that memory is searched in parallel, the serial processes performed by the computer that mimic this parallel search require more time when more entries must be compared. Hence, we minimized the number of pictures stored in the pattern activation subsystem and the number of objects described in associative memory. Because stimuli fall into classes, and processing is qualitatively the same within each class, for present purposes it made sense only to examine one member of each class. However, in principle an arbitrarily large number of pictures can be added to the pattern activation subsystem, with a corresponding increase in associative memory data structures (an earlier version of the program operated on 64 images).

### Effects of damage

The simulation is intended to allow us to predict new neuropsychological syndromes. Before the individual tasks are performed, the simulation can be damaged by selectively disrupting processing in individual subsystems or by disconnecting subsystems. According to the theory, such damage leads to difficulties in object identification because of four factors:

#### *Subsystems*

Each of the subsystems described above can be damaged, either completely or partially, resulting in a complete or partial failure to carry out the corresponding computation. In our computer simulations we have chosen only a subset of the possible types of partial damage, as will be discussed shortly. We use types of partial damage that seem plausible, given clinical case reports in the literature.

#### *Connections*

Any of the connections between subsystems can be disrupted. The consequence of disconnections is to cut off a subsystem's input or output (cf. Geschwind, 1965).

#### *Compensations*

There often are competing subsystems or memory representations in the system, and the normal balance among them can be altered following damage. At first glance, this possibility raises the

specter of hopelessly complex overdetermination of any behavioral deficit. However, according to the theory, subsystems do not assume qualitatively different functions following damage; compensations operate only by causing subsystems to be used in different circumstances following damage. Further, according to the theory even this limited form of compensation is highly constrained, and will in fact affect performance only in relatively complex tasks. According to our particular theory, such compensations can occur in three circumstances:

First, there often are numerous representations in associative memory that correspond to an input from the ventral system. The property lookup subsystems access all in parallel, and the one with the greatest "strength" (i.e., previously seen most often) "wins" (i.e., is used to guide top-down search). If this representation is damaged, the next strongest one is then used. This can result in a switch from using coordinate information to using categorical information..

Second, one purpose of the attention window is to yoke together object property and spatial property representations being processed in parallel in the inferior temporal lobe and parietal lobes, respectively. However, this mechanism alone is inadequate, if only because processing times are not necessarily the same in the two systems (and so it may be possible to encode several locations in the period required to encode a single shape or vice versa). Thus, we posit that upon receiving an input from the shape system, associative memory waits for an input from the location system before accepting another shape (it does not wait indefinitely, however, otherwise the entire system will freeze up following complete damage to the dorsal system). One consequence of this mechanism, in theory, is that damage to either system results in slowed down input from the other system.

Finally, output from the feature detection subsystem is very "fragile," varying depending on angle of regard, distance, lighting and so on. Thus, information via the preprocessing subsystem, which is relatively invariant over such vagaries, is weighted more strongly by associative memory. If the preprocessing or pattern activation subsystems, or any of their connections, are damaged, then the output from the feature detection will be used. Thus, in these cases same/different judgments can be carried out, but they will become very sensitive to perturbations in viewpoint and lighting.

#### *Activation*

Finally, damage can result in a decrease in "activation" level. In our model, the only consequence of this is a faster decay from short-term associative memory. That is, the input to associative memory is not stored long enough to build up very complex data structures, and hence identification that depends on multiple cycles through the system becomes difficult.

#### Generating Predictions

Forty-four distinct types of damage can be inflicted on the system. Not only can each connection be disrupted, but each subsystem can be completely or partially impaired. Complete damage to a subsystem or to the data lines leading to it have the effect of eliminating processing in that subsystem, so that no output is produced. Severing the output line(s) of a subsystem will not, of course, affect the

functioning of that particular subsystem, but because the output line from one subsystem is the input line to one or more connecting subsystems, the connecting subsystems will be deprived of their input, and this can affect additional subsystems further downstream depending on their information dependencies. Partial damage, on the other hand, alters the functions of subsystems so that their output differs from normal output in the ways indicated in Table 2.

-----  
 Insert Table 2 About Here  
 -----

With 44 individual types of possible damage, there are trillions of possible combinations of damage. This is a very daunting statistic both for users interested in exhaustively testing dysfunction hypotheses and for those responsible for assuring smoothly running program code! Fortunately, many of the combinations produce exactly the same kind of behavior, largely because damage occurring upstream frequently causes the system to fail before downstream connections and subsystems are even brought into play. In this article we consider only the effects of isolated damage to individual subsystems or connections, but it should be noted that various emergent properties will occur with combinations of damage (e.g., of the feature detection subsystem plus the preprocessing subsystem).

Thus, we examined the performance of the program on all of the tasks described below with each of the 44 possible dysfunctions tested in isolation. As will be seen, summary table was prepared of the results. This summary simply indicates whether there was success or failure for each task.

#### Simulated tasks

The system is capable of performing four kinds of tasks on a variety of inputs, as are illustrated in Figure 4. We will describe the tasks, and then indicate which subsystems are recruited during the course of performing them with the different pictures. Except when processing the overflowing face (which would fill two entire screens), where a second array is used when mimicking the eye movement necessary to find a part, one array remains in the visual buffer during all tasks. The attention window at first surrounds the whole image, and then encompasses different parts as necessary. The results of the simulations will be discussed after each task is described; further discussion will be deferred until we turn to the observed clinical syndromes.

-----  
 Insert Figure 4 About Here  
 -----

#### *Object class identification*

In our computer simulation, the user can select the question, "What is this?," and the system will attempt to produce an "entry level" name for the class of objects. That is, Jolicoeur, Gluck, and Kosslyn (1984) found that objects are named spontaneously at Rosch's (1978) "basic level" unless they are atypical of that category (e.g., a penguin for the category "bird"), in which case they are named at

a subordinate level. When subjects were asked to verify a name subordinate to the entry level (e.g., "canary," for which the entry level is "bird"), brief presentation of the picture followed by a mask had devastating effects – as expected, if top-down processing is necessary to search for additional distinctive properties, and the mask impaired this process. In contrast, if subjects were asked to verify a name superordinate to the entry level name (e.g., "bird" for penguin), brief presentation of the picture followed by a mask had relatively minor effects – as expected if this task requires inference in associative memory, and does not require collecting additional perceptual information.

Our computer simulation can perform this task with all eight stimulus pictures. This is of interest because processing is different in the different cases, as noted below:

-----  
Insert Table 3 About Here  
-----

*Familiar fox.* Table 3 presents the order in which subsystems are used when one names a familiar picture of a prototypical fox; subsystems entered with the same number are executed in parallel. Table 3 also provides a brief summary of processing at each step in this task. We assume that the program has seen this particular familiar picture of a fox many times, and so has a coordinate representation of the location of the parts as well as the categorical relations among them. According to our theory, and as is embodied in the simulation program, the overall fox pattern is similar enough to other patterns (e.g., a dog) that the match of trigger features in the pattern activation subsystem is not sufficient for confident identification, nor is the match of the pattern itself. Thus, top-down hypothesis testing is initiated. In this case, because the picture is very familiar, coordinate representations of spatial relations are very strong and hence are used over the corresponding categorical representations, and the attention window is moved to the location of a distinctive part (the head, in this case). As is schematized in Table 3 (section 7 and following), a second bottom-up cycle commences, encoding the part. The part does match the sought part (in the pattern activation subsystem), and the location of the part matches the sought location (in associative memory; in this case, because coordinate information is sought, the output from the coordinate relations encoding subsystem is used). This additional information is adequate for the identification threshold to be exceeded in associative memory.

-----  
Insert Table 4 About Here  
-----

Table 4 presents a summary of the results of running the simulation in the different tasks. The X marks in the task columns indicate when the program could not perform tasks, and the rows indicate the types of damage that engendered such failure. As is evident in Table 4, we found that the patterns of success and failure fell into fourteen categories, which are ordered from most severely disruptive to

least severely disruptive. The entries in the first column indicate that identifying a familiar fox is disrupted whenever the preprocessing, pattern activation, spatiotopic mapping, or coordinate relations encoding subsystems or their connections are damaged. Similarly, damaging associative memory or its inputs disrupts performance. Note that damaging the categorical property lookup subsystems did not disrupt task performance (IX and XII), but damaging the coordinate property lookup and coordinate encoding subsystems did (IV and VI). At first blush, this is puzzling, given that when the coordinate property lookup subsystem was disrupted, the categorical property lookup subsystem took over. However, the categorical-coordinate conversion subsystem needs coordinate information, which is accessed via the coordinate property lookup subsystem. Thus, damaging the coordinate property lookup subsystem has devastating effects when top-down processing is needed for any naming task. These results may seem counterintuitive, but recall the evidence that right-parietal lobe lesions affect picture naming (e.g., Warrington and Taylor, 1973). In contrast, when the connections from the coordinate property lookup subsystem to the attention shifting subsystem are disrupted, the system successfully compensated, using information about categorical relations to locate the sought part (VIII).

*Unfamiliar twisted fox.* This processing used to name an unfamiliar picture of a twisted fox differs from that described in Table 3 in three ways: First, because the picture is an unfamiliar contortion, there are no representations of the appropriate coordinate relations in associative memory, and hence categorical representations must be used to search for, and then check the locations of, distinctive parts (section 5 of Table 3). In this case, the program sought the head, which was located at the more tapered end of the general shape envelope. (Taper is not the only way to identify how an object is oriented, but it is relatively simple and thus was the heuristic we built into our model. Taper is encoded via the categorical relations encoding subsystem.) Second, when categorical relations are used, the categorical-coordinate conversion subsystem must also be used. This subsystem would be inserted between sections 5 and 6 of Table 3. Third, given that one is searching for a part in a categorically defined location, the output from the categorical relations encoding subsystem is used in associative memory to evaluate the input.

The interesting results here are those that distinguish this case from the first one. Note in Table 4 that damage to coordinate relations encoding and lookup disrupts both tasks because this information is needed by the categorical-coordinate conversion subsystem. However, partial damage to the coordinate relations encoding subsystem (resulting in the center of the image being miscalculated, case VI) did not affect processing in this task, whereas partial damage to the categorical property lookup subsystem that results in the wrong long-term memory structures being accessed, and related damage (case IX), did selectively disrupt performance on this task but not on the first one.

*Familiar occluded fox.* In this case, the trigger features and the overall image matched the stored representation of a fox rather poorly, and hence the output from the pattern activation subsystem was weak (indicating low confidence). Thus, the system needed additional evidence to evaluate this

hypothesis, and looked for two distinctive parts a fox should have. The program first found the head (the most distinctive part), and then tried to find the tail (the second most distinctive part). Unfortunately, the tail was occluded, and hence the program had to find the front legs (the third most distinctive part). Thus, the cycle from section 5 through section 10 in Table 3 was repeated three times in this case.

Note in Table 4 that damage affected this task exactly as it did the first one in all respects but one: partial damage to the coordinate lookup subsystem that causes it to perseverate will not allow enough different parts to be encoded to identify the object (XI). Perseveration is common following frontal lobe lesions, and it would be of interest to examine such a possible syndrome (cf. Luria, 1980).

*Unfamiliar twisted occluded fox.* The processing in this case was like that indicated in Table 3, except that now four cycles from section 5 to section 10 were necessary. Because the tail's location is specified relative to the body ("connected to the top rear side of the body"), and not the general shape envelope (as was the head's location), the program first had to locate the body before being able to search (unsuccessfully) for the tail.

As is evident in Table 4, processing disruptions here are identical to those for the twisted fox in all respects but one (XII): Now partial damage to the categorical property lookup subsystem that causes it to perseverate will not allow enough different parts to be encoded to identify the object. It would be of interest to examine a possible double dissociation, then, between the two types of perseveration and their predicted consequences on different types of object identification.

*Familiar face.* The familiar face is identified on the basis of matching trigger features in the pattern activation subsystem. Hence, only sections 1 through 4 of Table 3 are used, and the stored patterns themselves are not matched in the pattern activation subsystem (in section 3).

Because no top-down processing is used, performance is not affected by damage to any of the subsystems used to look up information and then to direct attention accordingly. Thus, we predict that some patients will be able to identify faces (as faces, not as particular people) but not be able to identify common objects. These results are interesting because they seem counterintuitive to many; they are a kind of "reverse prosopagnosia" (to be discussed below). Warrington (personal communication) reports having seen patients with this sort of selective deficit.

*Rotated familiar face.* The matching process in the pattern activation subsystem evaluates the consistency of trigger features in particular locations with those of a shape seen from a single point of view. Thus, for a highly familiar, distinctive and relatively rigid pattern such as a face, the system can identify the shape purely on the basis of matching trigger features even when the shape is rotated.

*Familiar occluded face.* Because faces are so distinctive, even the occluded face can be identified as a face on the basis of matching trigger features. Because faces are symmetrical, half of the face is sufficient to identify it as a face.

As is evident in Table 4, there was no selective effect of damage for the different variations of the face object-level identification task. This prediction contrasts strongly with the selective effects evident for the corresponding tasks with common objects, and hence would be easily tested.

*Overflowing familiar face.* Because faces are so distinctive, even the overflowing familiar face can be identified as a face on the basis of matching trigger features, as noted above.

#### *Exemplar identification*

In the computer simulation, one also can ask for an identification of a stimulus at the exemplar level. This task is performed only with the face stimuli, which was a matter of convenience: we expect that the same processing will occur whenever one is asked to identify a particular example of a stimulus class. We have used faces here mainly for historical reasons, given the early classification of prosopagnosia as being limited to faces. Thus, the user selects "Who is this?," and the simulation attempts to identify the particular face. The precise processing, however, depends on the particular type of stimulus.

Our key assumption here is that the face is sufficiently far from the viewer that its features are not represented at a high resolution. Thus, the trigger features extracted and the pattern itself will match those of many faces, and no particular hypothesis is formed. In this case, an entry level name ("face") is inappropriate, and hence is inhibited (as discussed above). Thus, a second pass will be necessary to encode individual features and their locations, allowing identification of the specific face.

*Familiar face.* Initial processing here is identical to that outlined in Table 3 in all respects but one: we assume that the question being posed has the result of inhibiting all but exemplar-level representations in the pattern activation subsystem and in associative memory; only the name of a specific exemplar can be used to answer the question appropriately. As is illustrated in Table 3 (section 3), after the trigger features are matched, the pattern itself is matched; the trigger features alone will not distinguish among faces. However, because of the scope/resolution tradeoff, the input pattern will not be sufficient to discriminate confidently among similar stored face patterns. Thus, the top-down hypothesis testing cycle will commence. In this case, coordinate representations are critical because the precise positions of features often are themselves part and parcel of the distinguishing properties of a face. Unlike the familiar picture of a fox, which could be identified using categorical relation representations if the coordinate relations could not be accessed, coordinate relations are critical here. Thus, no compensation occurs following damage to the connection between the coordinate property lookup subsystem and the attention shifting subsystem (as is evident in case VIII).

Thus, as is evident in Table 4, all damage that disrupted the use of coordinate information disrupted processing. It is of interest to compare these results to those with the fox pictures. For faces, any disruption of coordinate encoding disrupted performance, including the disconnection of the coordinate property lookup subsystem and the attention shifting subsystem; for foxes, the categorical property lookup subsystem compensated, and this information was used to direct attention.



*Rotated familiar face.* Because skulls are rigid objects, a face cannot be twisted in the same way as can a fox's body. Thus, we examined the effects of rotating a face 90°. Farah and Hammond (1988) have shown that a patient with a deficit in mental image rotation nevertheless could identify rotated pictures, which suggests that mental rotation is not used here. Indeed, Perrett et al. (1985) found that face-sensitive cells in STS often responded equally well to upright pictures and those rotated 90°, which is as expected if the same trigger features are extracted and the viewpoint consistency constraint is satisfied (Lowe, 1987b). Thus, the theory posits that processing here is identical to the previous case.

*Occluded familiar face.* Processing here is identical to the previous case, except that it may be necessary to encode more than a single part if the image is sufficiently degraded.

*Overflowing familiar face.* Processing here is identical to the previous case, except that multiple eye movements are necessary to take in the entire picture. Thus, note that partial damage to the attention shifting subsystem disrupted this task (case XIII), but not the others.

#### *Same/different discrimination*

The third type of task does not involve identifying pictures. In this case, the system was shown two pictures in succession (not simultaneously) and asked whether they are the same. The system performs this task in two ways. In one, it encodes each object by decomposing it into parts and coordinate relations among them, and then compares the corresponding representations of the two objects in associative memory. In the other, it uses the feature detection subsystem to encode the intensity gradients across the objects, and then compares the two representations in associative memory. The representations of intensity are given less weight if there is a conflict between the two procedures, because these representations are sensitive to local vagaries of lighting. Performance on the system was the same with different combinations of the different pictures being used as stimuli. In Table 4 we have presented the results when the stimuli consisted of a face and a fox.

Thus, we note in Table 4 that disrupting the visual buffer and associative memory disrupts this task (I), as does partial damage to associative memory that results in a very brief short-term memory (III). No other damage in isolation induced this deficit.

#### *Identifying pairs of pictures*

Finally, the simulation was shown two pictures side by side and asked "What is here?". In this case, it was to name both stimuli at an object level. This task is performed in normal processing by first adjusting the attention window to surround both pictures. The input to the ventral system does not match a stored representation (because the resolution is too poor to encode enough of the trigger features to permit a good match), but the dorsal system registers that there are two large perceptual units present. This information is stored in associative memory. Following this, the system chooses one (the left, in our model), and processes it just like a single object. Once this picture is identified, the system shifts to the other object and identifies it.

*Two familiar foxes.* In one version of the task, two foxes are presented. In this case, damage that disrupts identification of multipart objects in isolation also disrupts processing here. Furthermore, as is evident in Table 4, one form of partial damage of the spatiotopic mapping subsystem will disrupt this task but not the single-object identification tasks (X). In this case, partial damage causes the spatiotopic map to assign the same location (directly ahead, in our simulation) to all input. Thus, when it encodes a pair of objects, it registers only a single object (in the center of mass). It then attempts to identify the input by shifting attention to the left side of the object, looking for a part using a default bottom-up strategy that is adopted when no particular hypothesis is formed. It then encompasses the object on the left and processing proceeds apace. Once it has an hypothesis, attention is shifted top-down, using stored information, and the encoding of location is not critical (although it does require more cycles to confirm an hypothesis, using part matches without location matches). However, once it has finished with the object on the left, it does not "know" that there is a second object (only one location is represented in the spatiotopic map), and hence stays fixated on the left object until some other factor draws attention away.

*Two familiar faces.* The task again is to identify what is present at an object level. It is not surprising that identifying two faces is disrupted by damage that disrupts identifying a single face at the object level (cases I, II). In addition, damage that disrupted holding information in short-term memory (case III), encoding location (case IV), and looking up coordinates to direct attention (case IV and VIII) prevented the system from encoding two or more shapes present at the same time in the visual buffer. In all of these cases one or more form of single-object identification was also disrupted. Some types of damage, however, disrupted naming two foxes but not two faces at an object level (cases V, VI, and VII); in these cases additional parts needed to be encoded to name each fox (during hypothesis testing), but could not be matched effectively to stored information. Finally, this task alone was impaired when the spatiotopic mapping subsystem assigned only one location value to any given input—and hence the system was "unaware" that more than one object was present (case X). (Note the contrast to the tasks in which previously stored information was used to direct attention top-down, which did allow more than one part to be encoded.)

### III. INTERPRETING CLINICAL SYNDROMES

Although deficits in visual processing following brain damage have been described at least since the time of Hughlings Jackson, there have been controversies not only over the nature of such dysfunctions, but also over their very existence (e.g., see Bay, 1952). Furthermore, there is no universally accepted taxonomy for the various types of dysfunctions, which in part reflects the degree to which they are not well understood. Indeed, the notion that each type of dysfunction can be produced in many distinct ways belies the entire idea that a simple taxonomy will be very illuminating. The description of syndromes used here was derived primarily from Damasio (1986), DeRenzi (1982), and Williams (1970).

We will group the types of visual deficits following cortical damage into two classes. In the first class are deficits in the ability to represent and interpret perceptual units. These units correspond to shapes and parts thereof. In the second class are deficits in the ability to represent and interpret spatial relations among units. At the end of each section, we will relate the syndrome to our computer simulation results.

### Disruptions of the Representation of Object Properties

We begin by considering disorders of the representation of object properties, such as shape and color. "Visual agnosia" is a term used to cover a wide range of such perceptual deficits. Visual agnosia occurs when there is an impairment in recognition and identification that is not due to blindness per se, difficulty in naming, disruptions of attention, or general mental deterioration. Rather, the impairment is in the interpretation of the perceptual input (see Bauer and Rubens, 1984; Damasio, 1986; Humphreys and Riddoch, 1987, 1988; Levine, 1982; Ratcliff, 1982; Riddoch and Humphreys, 1987). This syndrome has also been called "mind blindness" or "psychic blindness," and was first demonstrated by Munk in 1881. Munk removed part of the occipital lobes from both hemispheres in a dog, and found that the animal could avoid bumping into objects but could not identify them. (The term "agnosia" was originally coined by Freud [1891], and came to be the accepted name for the syndromes.) These disorders are often seen along with lower-level disorders, such as blindness over one half of the visual field (homonymous hemianopia) or blind spots (scotoma). But they can occur in the absence of such lower-level disorders, and do not appear to be caused by them. Furthermore, they are usually specific to a given sensory modality; a person who cannot recognize or identify an object visually may be able to do so by touch or sound.

These disorders occur primarily following damage to the temporal lobes or damage to the occipito-temporo-parieto junction area, which presumably disconnects pathways running from the occipital lobe to the temporal lobe. Thus, the damage parallels the locations of lesions that produce the analogous deficit in monkeys (Levine, 1982), and can be conceptualized as damage to the ventral system, as described below.

### Visual object agnosia

Critchley (1953) describes a good example of a patient suffering from visual object agnosia (first reported by Bay, 1952), as follows:

"A sixty-year old man, almost blind in his right eye from an old injury, woke from a sleep unable to find his clothes, though they lay ready for him close by. As soon as his wife put the garments into his hands, he recognized them, dressed himself correctly and went out. In the streets he found he could not recognize people – not even his own daughter. He could see things, but not tell what they were." (page 289).

The typical description of visual object agnosia focuses on the patient's inability to recognize or identify objects he or she sees, but in the presence of a preserved ability to recognize and identify objects by their touch or sound. Two types of visual object agnosia are discussed: *Apperceptive* agnosia corresponds to difficulties in processing the sensory input and in putting together visual information gleaned over time (resulting in a conscious perception of the object); these patients cannot assess whether two objects are the same or different, let alone recognize or identify the objects. In contrast, *associative* agnosia corresponds to difficulties in making the connection between the perceptual input and previously stored information (e.g., De Renzi, 1982; Hecaen and Albert, 1979; Kolb and Whishaw, 1985). Thus, difficulties in appreciating the shape of an object are apperceptive; difficulties in identifying the object while still being able to distinguish its shape are associative. Patients with "pure" associative agnosia can discriminate shapes and make correct matches among shapes that are placed before them (in the pattern activation subsystem, according to the present theory), even though they cannot identify the shapes. These patients may be able to recognize an object (e.g., show evidence that it is familiar, such as taking longer to evaluate it) but not be able to identify it. The fact that the deficit is often modality-specific indicates that associative memory is intact, as are the mechanisms that use information after it is accessed.

The first eight columns of Table 4 correspond to tasks in which the simulation was asked, "What is this?". Wherever there is an X mark in one or more of these columns, but no X mark for the same/different task, we have an instance of behavior corresponding to that found with visual object agnosia. Thus, it is of interest that depending on the precise task, different sorts of damage affect processing. For common objects, 33 different types of damage produced some form of the deficit. In short, the theory predicts that "visual object associative agnosia" is not a single deficit, given that we expect dissociations between different types of stimuli in a simple object identification task and that we expect numerous types of damage to produce the impaired behavior.

#### Prosopagnosia

Prosopagnosia is a rare subclass of associative object agnosia; initially, it was thought to be limited solely to faces (hence the term "prosopon," meaning face). The clinical literature is replete with tales of patients who could identify everything but faces, including those of wives, children, siblings, and even themselves. There are documented cases in which a patient could not recognize or identify any of the people in a photograph that included the patient along with his friends (Williams, 1970, page 59). Indeed, there is a case reported (see Bauer and Rubens, 1985) in which a patient bumped into a mirror and apologized, apparently thinking it was another person (see also Humphreys and Riddoch, 1987). These patients typically can describe individual features but cannot put them together to identify a face. Recently it has been discovered that prosopagnosics are not impaired solely at face recognition and identification. Rather, they have difficulty in recognizing and identifying individual examples of objects that are identified by rather subtle variations in shape. So,

for example, a farmer could not identify his cows after he became prosopagnosic; a dog expert could not tell a potential prize winning purebred from an inferior animal; a bird expert had trouble picking out different birds, and so on (see Bruyer, 1986; Damasio, 1986; Hay and Young, 1982). Bilateral damage to the occipitotemporal junction area or mesial posterior inferior temporal lobes is the most usual correlate of prosopagnosia (Damasio, Damasio, and Van Hoesen, 1982).

Prosopagnosia is particularly intriguing in light of recent findings reported by Bauer (1984) and Tranel and Damasio (1985). These researchers found that patients with prosopagnosia showed marked electrodermal skin conductance responses (SKR) to previously seen faces, even when they could not identify them as being familiar. These responses were significantly larger than those to novel faces. Thus, at some level in the brain knowledge of familiar faces is registered.

Only the tasks that required the program to name the specific face bear on this deficit. No fewer than 24 types of damage produced this deficit. Of these, there were 16 types of damage that caused the system to be able to name the face as a face but not as an instance (cases III - VI and VIII). In all of these cases, the pattern activation subsystem is still able to recognize the shape. It is possible that some stored patterns (e.g., of specific parts) match the input better than others, but not enough better for a high-confidence output. Thus, an SKR response might be evinced even though a patient has no awareness that a somewhat better match has been made. Thus, this deficit cannot be considered a single entity. The only interesting difference among the face stimuli was the sensitivity of the system to partial damage of the attention shifting subsystem when an overflowing face was used (case XIII).

#### Color agnosia

The neurologist Holmes described a very selective deficit that beset an artist following a stroke (as described in Critchley, 1953, page 276), who was not able to use colors after the injury, as follows:

"He was not colour blind, however, for he could name most colours and pick out colours correctly to command. He could not associate colours with objects except by reference to rote memory."

Thus, this patient was not color blind, and had no difficulty in naming (although reading difficulties often occur along with color agnosia, and prosopagnosia is often accompanied by color agnosia; e.g., see Bruyer, 1986; Damasio, 1986; De Renzi, 1982). Unless the object-color association is verbally encoded by rote, these patients have difficulty recalling which colors belong to objects; for example, if given a black and white drawing of a strawberry, they are unable to select the correct crayon to color it properly.

The present theory has not been extended to provide accounts for this syndrome in detail. However, it is of interest that this sort of task typically requires imagery (Kosslyn, 1980). Indeed, the simplest account of the deficit is that the process that activates stored visual memories is awry. Kosslyn (1987) argues that the present theory can be extended in a simple way to account for imagery:

the pattern activation subsystem produces a pattern in the visual buffer, with patterns corresponding to parts being placed by moving attention just as is done during top-down hypothesis testing during perception. If so, then damage that leads to difficulties in top-down search should also lead to color agnosia, given that the color must be placed in specific parts of the shape.<sup>9</sup>

### Metamorphopsia

Although patients exhibiting this deficit may identify objects correctly, the percept is aberrant. "Macropsia" is a condition in which objects appear larger than they are, whereas "micropsia" is a condition in which they appear smaller than they are. Similarly, objects can appear fragmented, compressed, tilted, "turned around," and so on. One patient complained that faces looked like fish heads (see De Renzi, 1982). And the aberration is not necessarily static: one patient claimed that people's eyes would swell and contract, going from "nothing at all" and then coming back "like a pimple." In some of these patients object identification apparently is not dramatically disrupted.

This deficit would occur if the spatiotopic mapping subsystem did not compute size appropriately. If so, however, we predict that top-down search also will be disrupted.

### Disorders of the Representation of Spatial Relations

Shapes can correspond to objects or parts thereof. In either case, one often needs to be cognizant of the relative positions of the shapes – either the parts of a single object or the objects in a scene. The representation of position can be disrupted quite independently of the representation of shape. The major types of such disruptions are briefly described below.

### Simultanagnosia

The name of this disorder indicates what it is, namely an inability to grasp more than one shape at a time. For example, Williams (1970) describes the following patient, who had difficulty finding his way around because "he couldn't see properly", as follows:

It was found that if two objects (e.g., pencils) were held in front of him at the same time, he could only see one of them, whether they were held side by side, one above the other, or one behind the other.

Further testing showed that single stimuli representing objects or faces (including pictures) could be identified correctly and even recognized when shown again, whether simple or complex (newspaper photographs or simple sketches). If stimuli included more than one object, one only would be identified at a time, though the other would sometimes 'come into focus' as the first one went out...

If the patient was shown a page of drawings, the contents of which overlapped (i.e., objects drawn on top of one another), he tended to pick out a single object and deny that he could see any others. Moreover the figure selected at the first exposure of such stimuli was the only one seen on all subsequent presentations. If shown a drawing which might be seen in two different ways,

and which to the normal person usually appears first in one configuration and then in the other (reversible figures), he would pick out one configuration only and was quite unable to reverse it. (pages 62 - 63)

This syndrome was first characterized by Wolpert (1924), and was studied in detail by Luria (1959, 1973). The deficit is seen not only with separate objects, but also sometimes is found with the parts of a single object. Consider a case described by Tyler (1968): this patient "could see only one object or part of one object at a time...She reported seeing bits and fragments. For instance, when shown a picture of a U.S. flag, she said 'I see a lot of lines. Now I see some stars.'" Similarly, Goldenberg (personal communication) describes a patient who was asked to name a saw, and said "I see a round thing over here, and a line, and a jagged edge; must be a saw" and, when shown a picture of lion, "Looks like a cat's head. Here's a tail. Here's a tuft. Must be a lion."

Kinsbourne and Warrington (1962, 1963) showed pairs of stimuli to patients with simultanagnosia, and found that they had great difficulty reading both members of a pair. Perhaps more interesting, they showed these patients the stimuli successively, and measured the time required to identify them. They found that these patients could identify the first word of a sequence in a roughly normal amount of time, but had great difficulty with the second word. This effect diminished as more time was allowed between the presentation of the first and second words.

These sorts of disorders sometimes arise from bilateral damage to the occipito-temporo-parieto junction. In addition, there is a hint that the right parietal lobe may be more important in producing this disorder, but this inference must be viewed with caution (see De Renzi, 1982).

The performance of the simulation model when two objects were present at once illustrates this syndrome when the spatiotopic mapping subsystem was partially damaged, assigning the same location to all stimuli. Perhaps what is most interesting about this deficit is that, in its purest case when single multipart objects could still be identified, only one form of damage caused it. This result predicts that the deficit should be rare (relative to the other disorders), which it is.

#### Visuospatial disorientation

Other patients will visually mislocalize objects in space while still being able to identify them. Difficulty in localizing stimuli, usually diagnosed by difficulty in visually guided reaching (optic ataxia), without difficulty in identifying stimuli, is one component of a rare disorder called Balint's syndrome (see De Renzi, 1982; Damasio, 1986). (Two other components of this syndrome are described in the previous and following sections; however, there is controversy about the degree to which these deficits occur as a syndrome.) The syndrome was first explored in detail by Holmes (1919), who reported patients who could not use vision to guide reaching for objects, direct their gaze towards objects, estimate distance or navigate correctly. Some of these patients would bump into objects because they could not tell where they were relative to their bodies.

Localization difficulties most often arise following damage to the occipito-parietal area, typically in both hemispheres. There is a suggestion that the right hemisphere may be critical here, but this observation must be regarded as speculative. As is evident in Table 4 (in cases where top-down hypothesis testing is used), damage to the spatiotopic mapping, spatial relations encoding, attention shifting, property lookup subsystems or relevant connections can result in localization deficits. These deficits will also disrupt processing when multiple parts must be found during object identification. According to our simulation results (cases IV, VI, and VIII), intact object identification accompanied by visual mislocalization (assuming that motor control processes are intact) can only occur when single objects are identified purely on the basis of processing in the ventral system (without the necessity of encoding individual parts and relations). Only in case X will we find patients who can identify objects under normal circumstances, at both object and exemplar levels, but will fail to identify two adjacent objects that require multiple attention fixations. This form of damage disrupts a subsystem hypothesized to be implemented in the location most often associated with the actual lesions.

To our knowledge, such patients have never been tested carefully enough (e.g., by asking them to identify objects contorted in unusual ways, or subtending very large visual angles) to discover whether visual localization difficulty is often accompanied by difficulty in identifying multipart objects when top-down hypothesis testing is used, although Damasio (1986, p 278) notes that it is not uncommon for such patients to be misdiagnosed as having a form of agnosia (which can include a reported deficit in recognizing or identifying faces).

Furthermore, although localization difficulties often may be associated with simultanagnosia, Damasio (1986) reports having seen patients with simultanagnosia but without optic ataxia; these patients can point to objects they cannot recognize. Such a deficit will occur in our model only following damage to the visual buffer that degrades the input to the point where a perceptual unit can be registered (by the dorsal system) but no shape recognition or identification is possible. Consistent with this view, Damasio (1986) reports that such patients have bilateral damage restricted to the supracalcarine cortex.

#### Disorders of visual search

Some patients experience "paralysis of gaze" or visual scanning disorders (also called ocular apraxia) following brain damage (see Damasio, 1986; DeRenzi, 1982). These patients typically fixate on a stimulus, and cannot release their attention to look at another stimulus. As noted above, Posner and his colleagues (e.g., Posner et al., 1987) have identified three processes that may be involved in this kind of disorder. A patient may be unable to disengage attention from the previous stimulus; this deficit appears to be correlated with damage to the parietal lobes. He or she may be unable to shift attention (scan) to the next stimulus; this deficit appears to be correlated with damage to the superior colliculus. And a patient may be unable to engage attention once focused on the new stimulus; this deficit appears to be correlated with damage to the thalamus.



We have only roughly modeled attentional processes in our model. Deficits in these processes are the obvious source of a disorder of visual search, and thus it is of interest that this disorder also arises in our simulation in a variety of ways. First, it can occur due to damage to the property lookup subsystems, causing them to perseverate, looking up the same information repeatedly. Second, it can occur if the spatiotopic mapping subsystem is partially damaged, causing the system to assume that all objects are in the same place (directly in front, in our model). Third, damage to the attention shifting subsystem – even when it is delineated as coarsely as we have done – can result in this disorder. Fourth, damage to the ventral system that slows down object processing can cause a variety of this disorder if the dorsal system compensates. In this case, the attention shifting subsystem would fixate on an object until a useful input arrives at associative memory from the pattern activation or feature detection subsystems.<sup>10</sup>

#### Unilateral visual neglect and hemi-inattention

Some patients will ignore everything to one side of space, typically the left side (following right-hemisphere damage). If asked to copy a flower, they will draw only the petals on the right side of the plant; if asked to copy a clock, they either cram all of the digits into the right side or simply delete those that occur on the left side. If asked to bisect a line, these patients will place the bisector too far towards the right, as if the left side of the line were not present. These patients are not blind; they often can be led to pay attention to the neglected side, but only with great difficulty. The central nature of this phenomenon was demonstrated convincingly by Bisiach and Luzzatti (1977) and Bisiach, Luzzatti, and Perani (1978), who found it in visual mental imagery as well as in perception.

Unilateral visual neglect is most commonly observed during the acute stage of a patient's illness, in the first six weeks or so following the injury. There often is a course of recovery that begins with neglect and moves to "extinction with double simultaneous stimulation." In this middle part of the disease, a patient can see a stimulus to the left or to the right, but cannot see them both at once. This syndrome is different from simultanagnosia in that the simultanagnosic cannot see two overlapping stimuli in any position in the visual field, whereas if the neglect patient can see one stimulus (in the good field), he will be able to see the other.

One of the most fascinating aspects of the neglect syndrome is that the patients often do not realize that they have the disorder. Patients who have neglect with "anosognosia" do not compensate by moving their heads around or the like, and will deny that they have a problem. Indeed, one anecdote (heard by one of us at a hospital) describes a patient as worrying that she was going crazy because she "kept hearing voices" – namely those of people standing on her neglected side. She could not see them, and was not aware of her visual deficit. So she thought the voices were coming out of nowhere! This disorder hinges on a disruption of conscious experience, which is outside the domain of the present theory.

The neglect syndrome is very complex (e.g., see Heilman, Watson and Valenstein, 1985), and can arise following damage to at least four different locations in the brain (see Mesulam, 1981). The most common form of neglect arises from damage to the parietal lobe, particularly the right parietal lobe. It is clear that an understanding of neglect depends on details of the attention shifting subsystem that have not been developed here, and hence we will not attempt to address these phenomena. It is important to be aware of neglect, however, if only to distinguish between it and the syndromes described above.

#### IV. CONCLUSIONS

We had two goals at the outset of this article. Not only did we want to develop a theory of the component subsystems of high-level vision, but we wanted to use this theory to illuminate the causes of behavioral dysfunction following brain damage. The theory we developed was motivated by considerations of the abilities of the visual system, the neuroanatomy and neurophysiology of the visual system, and by analyses of the sort of information processing that is necessary to perform specific tasks. We have summarized experiments that tested key properties of the theory as we developed it. Only after we had a reasonably well-motivated theory did we implement a computer simulation model, and then generate predictions about the effects of brain damage. The interesting general result of our simulations is that there typically are many ways of obtaining the observed phenomena. In the future, then, neurological testing will have to be more subtle than is currently the norm (although see Humphreys and Riddoch, 1988, for several exceptions).

It is sometimes difficult to know what is important or distinctive about a theory as complex as the present one. Thus, it is worth underlining six critical distinctions the present theory offers for predicting object identification performance. To our knowledge, no other theory of similar detail emphasizes these distinctions. First, according to our theory a familiar shape can be identified in a single encoding if it is seen from a standard viewpoint and subtends no more than  $2^\circ$  of visual angle (so that its image falls on the fovea). In contrast, if a shape is not familiar, is in an unusual configuration, or subtends a large visual angle, then multiple encodings will be necessary to encode the separate parts and their spatial relations. Indeed, whenever an object subtends more than about  $2^\circ$  of visual angle, or is seen from an unfamiliar vantage point, then the initial match in the ventral system will be nonoptimal and the top-down hypothesis testing system will be called into play to encode distinctive parts and their locations. Marr (1982) does not address this issue, and Feldman (1985) would seem to predict that a single representation of an object can be built up in his "stable feature frame" regardless of the vantage point. Neither theorist makes a distinction between processing for familiar and unfamiliar shape configurations of an object.

The second key distinction arising from the theory is that when top-down search is necessary, spatial positions of parts and characteristics (e.g., a distinctive white spot on a cat's head) will be represented separately from the shape and object properties themselves. Ullman (1984) hints at a

similar distinction when he suggests that separate "visual routines" may be used to search for distinctive properties, but he makes no commitment to the way object property and spatial property information are stored. Feldman (1985) and Marr (1982), on the other hand, appear to posit a single representation that encompasses both shape and location.

The third key distinction is that two different kinds of spatial relation representations are encoded, and are useful in different circumstances. The present claim is that categorical spatial relations are useful when flexible objects are encoded, particularly when they are in unfamiliar configurations, whereas coordinate spatial relations are useful when relatively rigid objects that have subtle and important (for discriminating among similar objects) spatial relations are encoded. The distinction between categorical and coordinate representations may be implicit in Marr (1982), but is never explicitly developed. Similarly, Feldman (1985) does not address this distinction.

The fourth important distinction is that the method of top-down search will be different for rigid objects (such as a pencil, a quarter, or a baseball) and specific examples of objects that assume familiar shapes, on the one hand, and flexible objects seen that assume unfamiliar shapes (such as a sleeping dog, partially open scissors, or tumbled bicycle), on the other hand. In the former cases, coordinate representations will be used to direct attention to the presumed positions of parts, whereas in the latter cases, categorical representations will be used. Marr (1982) virtually ignores top-down processing in his theory, and Feldman (1985) does not posit separate processes in the two situations.

Fifth, to our knowledge no other computational theory has posited that there are both modality-specific long-term memories (such as the pattern activation subsystem posited here) and an amodal (propositional) memory (such as the associative memory posited here). Marr (1982) and Feldman (1985) both appear to posit exclusively amodal propositional memories.

Finally, perhaps the most distinctive feature of the present theory is its emphasis on formulating well-motivated hypotheses of distinct processing subsystems. Marr (1982) set the standard for how to engage in this kind of project, but limited himself primarily to low-level processes. Marr had very little to say about the "processing modules" (in his terminology) used in high-level vision. Feldman (1985) broke visual processing into four frames, with high-level processing being subsumed within an environment-centered frame and a "world knowledge formulary" (which essentially corresponds to a set of object-centered descriptions). To our knowledge, the present theory offers the first attempt to decompose high-level vision into computational processing subsystems. The subsystems posited by the present theory were formulated in light of constraints that do not permit a wide latitude of post-hoc or ad-hoc theorizing.

Our theorizing has been guided by what we call the "hierarchical decomposition constraint," the idea that smaller subsystems must be nested within those characterized at a coarser level. This requirement leads to a number of interesting possibilities. For example, the dorsal system may not, properly speaking, be part of the visual system per se. That is, it is possible that this system is

recruited in encoding location via other sensory modalities, such as audition or touch. Farah, Hammond, Levine and Calvanio (in press) present evidence for a nonvisual type of stored spatial representation, which could easily be generated by either the categorical or coordinate relations encoding subsystems.

Similarly, as noted above, the subsystems that work to shift attention to test hypotheses do not constitute a single subsystem at a coarser level of analysis, if we obey the hierarchical decomposition constraint. The property lookup subsystems presumably are also used in the service of language and other cognitive activities. In addition, the bottom-up processes that can shift attention (not represented here) presumably access the attention shifting subsystems directly, not in conjunction with the other subsystems discussed above.

Thus, there is some question as to how useful our intuitions will be in formulating theories about coarse subsystems, if we interpret subsystems as having direct mappings to neural activity. The traditional approach in cognitive science is to disavow this goal, and concentrate on abstracting regularities in stimulus/response relations. The theoretical entities typically are characterized purely at a functional level, with no thought to the complexity of the mapping to the underlying neural substrate. This approach seems appropriate for answering certain kinds of questions, such as identifying factors that must be respected when designing visual displays if humans are to use them effectively (e.g., Kosslyn, in press). However, if one's goal is to understand the effects of brain damage on behavior, then one must attempt to characterize what specific portions of the brain do. It may turn out that such functional descriptions are often counterintuitive, conflating what to commonsense seem like separate functions (such as color and shape in our preprocessing subsystem). Such characterizations should not be a surprise, given the way computation takes place in neural networks (for a relevant example, see Rueckl, Cave and Kosslyn, in press).

The present theory grew out of the initial effort described by Kosslyn (1987), which considered both imagery and perception. The same assumptions made there about common mechanisms also apply here. In particular, we assume that the pattern activation subsystem can produce a pattern in the visual buffer, which is the image proper. Once produced, the pattern in the buffer can be processed as it is in perception, encoding parts and characteristics and classifying them in various ways. In addition, we assume that multipart images can be built up by using the subsystems that shift attention to the locations of properties, only now images of the properties are placed after attention has been shifted to the proper locations (see Kosslyn, 1987). Farah (1988) reviews much evidence that imagery and like-modality perception share common neural mechanisms, and it would be straightforward to extend the present simulation to perform a number of imagery tasks – allowing us to generate predictions about the relationship between deficits of imagery and perception.

In this article we have focused almost exclusively on questions of *what* is accomplished by processing subsystems, not *how* it is accomplished in the subsystems. However, although it is tempting

to treat what and how as entirely distinct issues, they probably are intimately intertwined. Indeed, "what" at one level of analysis is really part of "how" at another; we have been decomposing object identification processing into ordered subsystems, which could be viewed as specifying "how" processing is done at a relatively coarse level. Although the distinction between what and how seems clear enough at a given level of analysis, in general it may be better to think about different levels of coarseness of an algorithm. This orientation seems reasonable in part because the two levels are mutually interdependent: what has to be computed depends in part on how computations are performed. For example, if each subsystem corresponds to a parallel distributed network, as we have assumed here, the problems that must be solved are different than if some subsystems are carried out using standard serial algorithms (where data structures are separate from processes). If networks are used, learning, searching and comparison have different properties than is typical in more traditional algorithms (see Rumelhart and McClelland, 1986).

The present effort has produced the outlines of major subsystems used in high-level vision. We have no doubt that each of these subsystems can be further decomposed into more specific information processing components. However, if we are correct we have delineated major classes of such subsystems, defining the general classes of processing that are performed. In this article we have assumed that each subsystem corresponds to a distinct neural network, but we have made no effort to implement such a complicated system using neural networks. It will be of interest to discover whether parallel distributed networks can carry out the processes we have hypothesized, and whether they can do so in real time and with appropriate levels of accuracy.

## Footnotes

Requests for reprints should be sent to S. M. Kosslyn, 1236 William James Hall, 33 Kirkland Street, Cambridge, MA 02138. This work was supported by ONR contract N00014-85-K-0291 and AFOSR grant 88-0012. We thank Jerome Feldman, John Gabrieli, Ron Katriel, Olivier Koenig, David Lowe, Vera Maljkovic, Michael Posner, Robert Rafal, V. S. Ramaschandran, Jay Rueckl, Shimon Ullman, David Van Essen, and especially David Mumford, for valuable comments and criticisms. We also thank Jack Chen and Steve Engel for skilled assistance in implementing an earlier version of the simulation model.

- 
1. Kosslyn (1987) treated these abilities as problems to be solved by the system, focusing on three specific abilities. Although this perspective does serve to emphasize the necessity of developing explicit information-processing mechanisms that are capable of producing the behavior, it has some unfortunately teleological overtones. Thus, in this article we consider the observed behavior of a system as the starting point, which needs to be explained by reference to underlying mechanisms.
  2. Kosslyn (1987) referred to the two pathways as processing shape versus location. We have changed our characterization to "object properties" versus "spatial properties" because color, properly speaking, is not an aspect of shape per se, and size and orientation are not location per se. It should not be surprising that it will prove difficult to find natural language terms that adequately characterize these computational systems.
  3. Gross and Mishkin motivate their hypothesis in part by reference to the very large receptive fields that characterize neurons in IT. Unfortunately, neurons in the parietal lobe also have very large receptive fields. A critical difference, perhaps, between the two is the observation that IT neurons typically contain the fovea, which produces the highest output, but it appears in different places within the receptive field. In contrast, parietal neurons often do not include the fovea. Thus, the response profiles of overlapping receptive fields will be different, with less systematic overlap in the IT neurons. This property may impair the use of "coarse coding" (Hinton, McClelland, and Rumelhart, 1986) to compute location in IT.
  4. Note that various other sorts of preprocessing must take place earlier in the information-processing stream, such as those involved in finding edges and growing regions in the input; the present preprocessing subsystem deals only with the preprocessing described here, marking trigger features prior to shape matching.
  5. The Lowe and Ullman schemes differ in numerous ways, particularly in the types of transformations of the representations that are allowed during comparison. Because we are focusing on the

nature of the processing subsystems, we will remain agnostic over the proper algorithm for matching input to stored shape representations.

6. The version of the theory presented in Kosslyn (1987) placed a greater emphasis on decomposing objects into parts. According to the present version of the theory, the entire object is processed initially (if viewing circumstances permit), and only if it fails to match a stored representation in the pattern activation subsystem are the parts encoded individually.
7. There are several changes in Table 1 compared to Table 1 of Kosslyn (1987). The only important differences are as follows: First, the notion of a "shape encoding subsystem" has been decomposed into the three subsystems of the ventral system (preprocessing, pattern activation, and feature detection). Not only was this single subsystem too coarsely characterized before to be useful for understanding the syndromes, but it was too coarsely characterized to allow us to implement it in a computer program. Furthermore, the name was unfortunate: According to the present theory, the ventral system encodes more than shape per se, and if multiple fixations are necessary both the ventral and dorsal systems are necessary to encode a shape. Second, the "coordinate location subsystem" has been renamed the "coordinate relations subsystem," and we no longer posit that only a single-origin representation is possible. This change came out of a deeper consideration of the information-processing tasks to be solved. For example, when deciding whether to put one's foot between two rocks while hiking, one wants to represent not only the distance between two rocks relative to each other but also the orientation of the rocks and the distance of the pair relative to oneself. Third, the "categorical relations access and interpretation" subsystem has been replaced by the present categorical property lookup and categorical-coordinate conversion subsystem. The motivation for this refinement (which is described in the text) became apparent as we considered how to implement the process. Finally, slight rewording has been used to clarify the operations performed by the other subsystems. When new subsystems were hypothesized, then, the hierarchical decomposition constraint was obeyed.
8. It is worth emphasizing that according to the present theory it is not necessary to segment the images into parts prior to entering the high-level visual system. We mark the separate parts in our simulations only for convenience. As noted above, this segmentation is useful during the top-down hypothesis-testing cycle when a specific part is sought. In this case, marking each part with a different letter allows the program to encode only the region of the pattern encompassed by the attention window that belongs to a single part. This initial filtering then simplifies the process of matching the encoded pattern to stored patterns. In theory, use of the "viewpoint consistency constraint" by the pattern activation subsystem obviates this virtue: As long as enough trigger features and their locations in the visual buffer are encoded from the sought

part, the input should be matched to a stored pattern—even if some additional trigger features from contiguous parts are also encoded at the same time (cf. Lowe, 1987a, b).

9. Color agnosia must be distinguished from acquired achromatopsia. The patient suffering from acquired achromatopsia can no longer perceive color (see De Renzi, 1982; Damasio, 1986). Acquired achromatopsia is highly correlated with prosopagnosia. It is possible that slight differences in shading are used to compute both shape and color (underlying the Land effect), and that disruptions of this process also slightly degrade the input to the ventral system, making subtle matches difficult.
10. The model described in this article does not assess processing time, and hence this compensation is not evinced in the behavior of this simulation.



## References

- Allman, J. M., and Kaas, J. H. (1976). Representation of the visual field on the medial wall of occipital-parietal cortex in the owl monkey. *Science*, 191, 572-575.
- Andersen, R. A., Essick, G. K., and Siegel, R. M. (1985). Encoding of spatial location by posterior parietal neurons. *Science*, 230, 456-458.
- Anderson, J. R. (1983). Spreading activation. In J. R. Anderson and S. M. Kosslyn (Eds.), *Tutorials in Learning and Memory: Essays in Honor of Gordon H. Bower*. San Francisco, CA: W. H. Freeman.
- Attneave, F. (1974). How do you know? *American Psychologist*, 29, 493-499.
- Bay, E. (1952). Der gegenwartige Stand der Aphasie-Forschung, *Folia Phoniatrica*, 4, 9-30.
- Bauer, R. M. (1984). Autonomic recognition of names and faces in prosopagnosia: a neuropsychological application of the guilty knowledge test. *Neuropsychologia*, 22, 457-469.
- Bauer, R. M., and Rubens, A. B. (1985). Agnosia. In K. M. Heilman and E. Valenstein (Eds.), *Clinical Neuropsychology*. New York: Oxford University Press.
- Benton, A. L. (1982). Spatial thinking in neurological patients: historical aspects. In M. Potegal (Ed.), *Spatial Abilities: Developmental and Physiological Foundations*. New York: Academic Press.
- Benton, A. L. (1985). Visuo-perceptual, visuo-spatial, and visuo-constructive disorders. In K. M. Heilman and E. Valenstein (Eds.), *Clinical Neuropsychology*. New York: Oxford University Press.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94, 115-147.
- Bisiach, E., and Luzzatti, C. (1978). Unilateral neglect of representational space. *Cortex*, 14, 129-133.
- Bisiach, E., Luzzatti, C., and Perani, D. (1979). Unilateral neglect, representational schema, and consciousness. *Brain*, 102, 609-618.
- Bower, G. H., and Glass, A. L. (1976). Structural units and the reintegrative power of picture fragments. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 456-466.
- Bower, T. (1970). *Perceptual Development*. San Francisco: W. H. Freeman.
- Bruyer, R. (1986). *The Neuropsychology of Face Perception and Facial Expression*. Hillsdale, NJ: Erlbaum.
- Bruce, C., Desimone, R., and Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*, 46, 369-384.
- Bundesen, C., and Larsen, A. (1975). Visual transformation of size. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 214-220.
- Campbell, F. W. (1980). The physics of visual perception. *Philosophical Transactions of the Royal Society of London*, B290, 5 - 9.
- Cave, C. B., and Kosslyn, S. M. (1988). The relative importance of parts and relations in object recognition. Harvard University manuscript.

- Cave, K. R., and Kosslyn, S. M. (in press). Varieties of size-specific selection. *Journal of Experimental Psychology: General*
- Colby, C. L., and Miller, E. K. (1986). Eye movement related responses of neurons in superior temporal polysensory area of macaque. *Neurosciences Abstracts*, 12, 1184.
- Cowey, A. (1985). Aspects of cortical organization related to selective attention and selective impairments of visual perception: a tutorial review. In M. S. Posner and O. S. Marin (Eds.), *Attention and Performance XI*. Hillsdale, NJ: Erlbaum.
- Critchley, M. (1953). *The Parietal Lobes*. Edward Arnold: London.
- Damasio, A. R. (1986). Disorders of complex visual processing: agnosias, achromatopsia, Balint's syndrome and related difficulties of orientation and construction. In M.-M. Mesulam (Ed.), *Principles of Behavioral Neurology*. Philadelphia, PA: F. A. Davis. (pp 259-288)
- Damasio, A. R., Damasio, H., and Van Hoesen, G. W. (1982). Prosopagnosia: Anatomic basis and behavioral mechanisms. *Neurology (NY)*, 32, 331-341.
- De Renzi, E. (1982). *Disorders of Space Exploration and Cognition*. New York: John Wiley & Sons.
- Desimone, R., Albright, T. D., Gross, C. G., and Bruce, C. J. (1984). Stimulus selective properties of inferior temporal neuron in the macaque. *Journal of Neuroscience*, 4, 2051-2062.
- DeYou, E. A., and Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in NeuroSciences*, 11, 219-226.
- Diamond, R., and Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, 115, 107-117.
- Downing, C. J., and Pinker, S. The spatial structure of visual attention. In M. I. Posner and O. S. M. Marin (Eds.) *Attention and Performance XI*. Hillsdale, NJ: Erlbaum. (pp. 177 - )
- Egeth, H. (1977). Attention and preattention. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* 11. New York: Academic Press (pp 277 - 320)
- Eriksen, C. W., and St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics*, 40, 225-240.
- Farah, M. J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review*, 95, 307 - 317.
- Farah, M. J., and Hammond, K. M. (1988). Mental rotation and orientation-invariant object recognition: dissociable processes. *Cognition*, 29, 29 - 46.
- Farah, M. J., Hammond, K. M., Levine, D. N., and Calvanio, R. (in press). Visual and spatial mental imagery: dissociable systems of representation. *Cognitive Psychology*
- Feldman, J. A. (1985). Four frames suffice: A provisional model of vision and space, *Behavioral and Brain Sciences*, 8, 265-289.

- Freud, S. (1953). *On Aphasia*. Trans. E. Stengel. New York: International Universities Press.  
(Originally published 1891)
- Geschwind, N. (1965). Disconnexion syndromes in animals and man. *Brain*, 237-294, 585-644.
- Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational knowledge. In F. Plum and V. Mountcastle (Eds.), *Handbook of Physiology*, Vol 5. Washington, D.C.: American Physiology Society. (pp. 373-417)
- Goodman, N. (1968). *Languages of Art*. Indianapolis: Bobbs-Merrill.
- Gregory, R. L. (1970). *The Intelligent Eye*. London: Weidenfeld and Nicholson.
- Gross, C. G. (1973). Visual functions of inferotemporal cortex. In R. Jung (Ed.), *Handbook of Sensory Physiology*, Vol 7. Berlin: Springer-Verlag.
- Gross, C. G. (1978). Inferior temporal lesions do not impair discrimination or rotated pattern in monkeys. *Journal of Comparative and Physiological Psychology*, 92, 1095-1109.
- Gross, C. G., Bruce, C. J., Desimone, R., Fleming, J., and Gattass, R. (1981). Cortical visual areas of the temporal lobe. In C. N. Woolsey (Ed), *Cortical Sensory Organization II: Multiple Visual Areas*. Clinton, NJ: Humana Press.
- Gross, C. G., Desimone, R., Albright, T. D., and Schwartz, E. L. (1984). Inferior cortex as a visual integration area. In F. Reinoso-Suarez and C. Ajmone-Marsan (Eds.), *Cortical Integration*. New York: Raven Press.
- Gross, C. G., and Mishkin, M. (1977). The neural basis of stimulus equivalence across retinal translation. In S. Harnad, R. Doty, J. Jaynes, L. Goldstein, and G. Krauthamer (Eds), *Lateralization in the Nervous System*. New York: Academic Press.
- Hannay, H. J., Varney, N. R., and Benton, A. L. (1976). Visual localization in patients with unilateral brain disease. *Journal of Neurology, Neurosurgery and Psychiatry*, 39, 307-313.
- Hay, D. C., and Young, A. W. (1982). The human face. In A. W. Ellis (Ed.), *Normality and Pathology in Cognitive Function*. New York: Academic Press.
- Hecaen, H., and Albert, M. L. (1978). *Human Neuropsychology*. New York: John Wiley and Sons.
- Heilman, K. M., and Valenstein, E. (Eds.). (1985). *Clinical Neuropsychology*. (Second Edition). New York: Oxford University Press.
- Heilman, K. M., Watson, R. T., and Valenstein, E. (1985). Neglect and related disorders. In K. M. Heilman and E. Valenstein (Eds.), *Clinical Neuropsychology* (Second Edition). New York: Oxford University Press.
- Hinton, G. E., McClelland, J. L., and Rumelhart, D. E. (1986). Distributed processing. In Rumelhart, D. E., and J. L. McClelland (Eds.). *Parallel Distributed Processing, Vols I*. Cambridge, MA: MIT Press. pp 77 - 109.
- Hinton, G. E., and Anderson, J. A. (1981). *Parallel Models of Associative Memory*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Hock, H., Kronseder, C., and Sissons, S. (1981). Hemispheric asymmetry: the effect of orientation on same different comparison. *Neuropsychologia*, 19, 723-727.
- Hoffman, D. D., and Richards, W. A. (1984). Parts of recognition, *Cognition*, 18, 65-96.
- Holmes, G. (1919). Disturbances of visual space perception. *British Medical Journal*, 2, 230-233.
- Holmes, E.J., and Gross, C. G. (1984). Effects of temporal lesions on discrimination of stimuli differing in orientation. *Journal of Neuroscience*, 4, 3063-3068.
- Hubel, D. H., and Livingstone, M. S. (1987). Segregation of form, color and stereopsis in primate area 18. *Journal of Neuroscience*, 7, 3378-3415.
- Humphreys, G. W., and Riddoch, M. J. (1987). *To See But Not To See: A Case Study of Visual Agnosia*. Hillsdale, NJ: Erlbaum.
- Humphreys, G. W., and Riddoch, M. J. (Eds.) (1988). *Visual object processing: A cognitive neuropsychological approach*. London: Erlbaum.
- Ingle, D. (1967). Two visual mechanisms underlying the behavior of fish. *Psychologische Forschung*, 31, 44-51.
- Jackson, J. H. (1864). Clinical remarks on defects of expression (by words, writing, signs, etc.) in diseases of the nervous system. *Lancet*, 1, 604-605.
- Jackson, J. H. (1874). On the duality of the brain. Reprinted (1932) in J. Taylor (Ed.), *Selected Writings of John Hughlings Jackson*. London: Hodder and Stoughton.
- Jolicoeur, P., and Kosslyn, S. M. (1983). Coordinate systems of visual long-term memory representations. *Cognitive Psychology*, 15, 301-345.
- Jolicoeur, P., Gluck, M. A., and Kosslyn, S. M. (1984). Words and pictures: Making the connection. *Cognitive Psychology*, 16, 243 - 275.
- Johnston, A. (1986). A spatial property of the retino-cortical mapping. *Spatial Vision*, 1, 319-331.
- Jonides, J. (1983). Further toward a model of the mind's eye's movement. *Bulletin of the Psychonomics Society*, 21, 247-250.
- Kinsbourne, M., and Warrington, E. K. (1962). A disorder of simultaneous form perception. *Brain*, 85, 461-486.
- Kinsbourne, M., and Warrington, E. K. (1963). The localizing significance of limited simultaneous visual form perception. *Brain*, 86, 697-702.
- Koenig, O. and Gabrieli, J. D. E. (1988). On the development of categorical spatial relations. Harvard University manuscript.
- Kolb, B. and Whishaw, I. Q. (1985). *Fundamentals of Human Neuropsychology (2nd Ed.)*. San Francisco: W. H. Freeman.
- Kosslyn, S. M. (1980). *Image and Mind*. Cambridge, MA: Harvard University Press.
- Kosslyn, S. M. (1983). *Ghosts in the Mind's Machine*. New York: W. W. Norton.

- Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres: a computational approach. *Psychological Review*, 94, 148-175.
- Kosslyn, S. M., Koenig, O., Barrett, A., Cave, C. B., Tang, J., and Gabrieli, J. D. E. (in press). Evidence for two types of spatial representations: Hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance*.
- Larsen, A., and Bundesen, C. (1978). Size scaling in visual pattern recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 1 - 20.
- Latto, A., Mumford, D., and Shah, J. (1984). The representation of shape. In *Proceedings of IEEE Workshop on Computer Vision, Annapolis, MD*. IEEE: Washington, D.C.
- Levine, D. N. (1982). Visual agnosia in monkey and man. In D. J. Ingle, M.A. Goodale, and R. J. W. Mansfield (Eds.), *Analysis of Visual Behavior*. Cambridge, MA: MIT Press.
- Livingstone, M. S., and Hubel, D. H. (1987a). Connections between layer 4B of area 17 and the thick cytochrome oxidase stripes of area 18 in the squirrel monkey. *Journal of Neuroscience*, 7, 3371-3377.
- Livingstone, M. S., and Hubel, D. H. (1987b). Psychophysical evidence for separate channels for the perception of form, color, movement and depth. *Journal of Neuroscience*, 7, 3416-3468.
- Lowe, D. G. (1987a). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, 31, 355-395.
- Lowe, D. G. (1987b). The viewpoint consistency constraint. *International Journal of Computer Vision*, 1, 57-72.
- Luria, A. R. (1959). Disorders of "simultaneous perception" in a case of bilateral occipito-parietal brain injury. *Brain*, 82, 437-449.
- Luria, A. R. (1980). *Higher Cortical Functions in Man*. New York: Basic Books.
- Lynch, J. C., Mountcastle, V. B., Talbot, W. H., and Yin, T. C. T. (1977). Parietal lobe mechanisms for directed visual attention. *Journal of Neurophysiology*, 40, 362-389.
- Marr, D. (1982). *Vision*. San Francisco, CA: W. H. Freeman.
- Maunsell, J. H. R., and Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, 10, 363-401.
- Mehta, Z., Newcombe, F., and Damasio, H. (1987). A left hemisphere contribution to visuospatial processing. *Cortex*, 23, 447-461.
- Mesulam, M-M. (1981). A cortical network for directed attention and unilateral neglect. *Annals of Neurology*, 10, 309-325.
- Milner, B. (1968). Visual recognition and recall after right temporal-lobe excision in man. *Neuropsychologia*, 6, 191-210.
- Mishkin, M. (1982). A memory system in the monkey. *Philosophical Transactions of the Royal Society of London B*, 298, 85-95.

- Mishkin, M., and Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research*, 6, 57-77.
- Mishkin, M., Ungerleider, L. G., and Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in NeuroSciences*, 6, 414-417.
- Miyashita, Y., and Chang, H. S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature*, 331, 68-70.
- Moran, J., and Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229, 782-784.
- Munk, H. (1881). *Über die Funktionen der Grosshirnrinde*. Berlin: Hirschwald.
- Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton- Century-Crofts.
- Neisser, U. (1976). *Cognition and Reality*. San Francisco: W. H. Freeman.
- Olson, D., and Bialystok, E. (1983). *Spatial Cognition: the Structure and Development of Mental Representations of Spatial Relations*. Hillsdale, NJ: Erlbaum.
- Palmer, S. E. (1977). Hierarchical structure in perceptual representations. *Cognitive Psychology*, 9, 441-474.
- Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., and Jeeves, M. A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society London, B* 223, 293-317.
- Pinker, S. (1985). Visual cognition: an introduction. In S. Pinker (Ed.), *Visual Cognition*. Cambridge, MA: MIT Press.
- Pinker, S., and Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73 - 194.
- Pohl, W. (1973). Dissociation of spatial discrimination deficits following frontal and parietal lesions in monkeys. *Journal of Comparative and Physiological Psychology*, 82, 227-239.
- Posner, M. I., Inhoff, A. W., Friedrich, F. J., and Cohen, A. (1987). Isolating attentional systems: A cognitive-anatomical analysis. *Psychobiology*, 15, 107 - 121.
- Posner, M. I., Snyder, C. R. R., and Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109, 160-174.
- Posner, M. I., Walker, J. A., Friedrich, F. J., and Rafal, R. D. (1984). Effects of parietal lobe injury on covert orienting of visual attention. *Journal of Neuroscience*, 4, 1863-1974.
- Rafal, R. D., and Posner, M. I. (1987). *Proceedings of the National Academy of Sciences USA*, 84, 7349 -
- Ratcliff, G. (1979). Spatial thought, mental rotation and the right cerebral hemisphere. *Neuropsychologia*, 17, 49-54.
- Ratcliff, G. (1982). Disturbances of spatial orientation associated with cerebral lesions. In M. Potegal (Ed.), *Spatial Abilities: Developmental and Physiological Foundations*. New York: Academic Press.

- Ratcliff, G., and Davies-Jones, G. A. G. (1972). Defective visual localisation in focal brain wounds. *Brain*, 95, 49-60.
- Reed, S. K., and Johnsen, J. A. (1975). Detection of parts in patterns and images. *Memory and Cognition*, 3, 569-575.
- Riddoch, M. J., and Humphreys, G. W. (1987). A case of integrative visual agnosia. *Brain*, 110, 1431-1462.
- Rizzolatti, G., Gentilucci, M., and Matelli, M. (1985). Selective spatial attention: One center, one circuit, or many circuits? In M. I. Posner and O. S. M. Marin (Eds.), *Attention and Performance XI*. Hillsdale, NJ: Erlbaum.
- Rosch, E. (1978). Principles of categorization. In E. Rosch and B. B. Lloyd (Eds.), *Cognition and Categorization*. Hillsdale, NJ: Erlbaum.
- Rueckl, J. G., Cave, K. R., and Kosslyn, S. M. (in press). Why are "what" and "where" processed by separate cortical systems? A computational investigation. *Journal of Cognitive Neuroscience*.
- Rumelhart, D. E., and J. L. McClelland (Eds.) (1986a). *Parallel Distributed Processing, Vols I and II*. Cambridge, MA: MIT Press.
- Rumelhart, D. E., and J. L. McClelland (1986b). On learning the past tense of English verbs. In J. L. McClelland and Rumelhart, D. E. (Eds.) (1986a). *Parallel Distributed Processing, Vols II*. Cambridge, MA: MIT Press. pp 216 - 271.
- Sagi, D., and Julesz, B. (1985). "Where" and "what" in vision. *Science*, 228, 1217-1219.
- Schiffman, H. R. (1982). *Sensation and Perception*. New York: John Wiley.
- Schneider, G. E. (1969). Two visual systems. *Science*, 163, 895-902.
- Shulman, G. L., and Wilson, J. (1987). Spatial frequency and selective attention to local and global information. *Perception* 16, 89-101.
- Schwartz, E. L., Desimone, R., Albright, T. D., and Gross, C. G. (1984). Shape recognition and inferior temporal neurons. *Proceedings of the National Academy of Science USA*, 80, 5776-5778.
- Smith, E. E., and Medin, D. (1983). *Categories and Concepts*. Cambridge, MA: Harvard University Press.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74, (11, Whole no. 498).
- Taylor, A. M., and Warrington, E. K. (1973). Visual discrimination in patients with localized brain lesions. *Cortex*, 9, 82-93.
- Tootell, R. B. H., Silverman, M. S., Switkes, E., and De Valois, R. L. (1982). Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science*, 218, 902-904.
- Tranel, D., and Damasio, A. R. (1985). Knowledge without awareness: an autonomic index of facial recognition by prosopagnosics. *Science*, 228, 1453-1454.

- Treisman, A. M., and Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- Tyler, H. R. (1968). Abnormalities of perception with defective eye movements (Balint's syndrome). *Cortex*, 4, 154-171.
- Ullman, S. (1984). Visual routines. *Cognition*, 18, 97-159.
- Ullman, S. (1986). An approach to object recognition: aligning pictorial descriptions. MIT AI Memo 931, December, 1986.
- Ungerleider, L. G., and Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield (Eds.), *Analysis of Visual Behavior*. Cambridge, MA: MIT Press.
- Van Essen, D. (1985). Functional organization of primate visual cortex. In A. Peters and E. G. Jones (Eds.), *Cerebral Cortex (Vol. 3)*. New York: Plenum Press.
- Van Essen, D. C., and Maunsell, J. H. R. (1983). Hierarchical organization and functional streams in visual cortex. *Trends in Neuroscience*, 6, 370-375.
- Warrington, E. K., James, M., and Kinsbourne, M. (1966). Drawing disability in relation to laterality of lesions. *Brain*, 89, 53-82.
- Warrington, E. K., and Rabin, P. (1970). Perceptual matching in patients with cerebral lesions. *Neuropsychologia*, 8, 475-487.
- Warrington, E. K., and Taylor, A. M. (1973). The contribution of the right parietal lobe to object recognition, *Cortex*, 9, 152-164.
- Weisenberg, T., and McBride, K. E. (1935). *Aphasia: A Clinical and Psychological Study*. New York: Commonwealth Fund.
- Williams, M. (1970). *Brain Damage and the Mind*. Penguin Book: Middlesex, England.
- Wilson, H. R., and Bergen, J. R. (1979). A four mechanism model for spatial vision. *Vision Research*, 19, 19-32.
- Wittgenstein, L. (1953). *Philosophical Investigations*. New York: Macmillan.
- Wolpert, I. (1924). Die Simultanagnosie. Störung der Gesamtauffassung. *Ztschr. Ges. Neurol. Psychiatr.*, 93, 397-415.
- Yarbus (1967). *Eye Movements and Vision*. New York: Plenum Press.
- Yin, T.C.T., and Mountcastle, V. B. (1977). Visual input to the visuomotor mechanisms of the monkey's parietal lobe. *Science*, 197, 1381-1383.



Table 1: Summary of Hypothesized Subsystems

SUBSYSTEM	INPUT	OPERATION	OUTPUT	LOCALIZATION (?)
<i>Spatiotopic mapping</i>	location of perceptual units in visual buffer, attention window, eyes, head, body	produces a representation of location of perceptual units in space	map of locations	posterior parietal lobes
<i>Categorical relations encoding</i>	locations of two perceptual units in space	computes categorical relation between units	categorical spatial relation	posterior left parietal lobe
<i>Coordinate relations encoding</i>	locations of two perceptual units in space	computes coordinates of one unit relative to another	coordinates	posterior right parietal lobe
<i>Preprocessing</i>	pattern from attention window	marks nonaccidental trigger features	trigger features marked on shape	occipito-temporal cortex
<i>Pattern activation</i>	trigger features marked on shape, feature encodings	matches trigger features and shape to stored pattern	shape identification code, plus goodness-of-match index	anterior inferior temporal cortex
<i>Feature detection</i>	patterns of activity in visual buffer	detects color, texture, and intensity	feature identification code	circumstriate occipital cortex
<i>Associative memory</i>	shape, feature identification codes, categorical and coordinate spatial relations	converges on object representation that is most consistent with input	information associated with object	posterior superior temporal cortex
<i>Categorical property lookup</i>	instruction to look up name and location of salient part	accesses name and location of salient part in associative memory until finds subgoals linking present location and to-be-sought location	name code sent to pattern activation subsystem and location information sent to categorical-coordinate-conversion subsystem	left prefrontal cortex
<i>Coordinate property lookup</i>	instruction to look up name and location of salient part	accesses name and location of salient part in associative memory	name code sent to pattern activation subsystem and location information sent to attention shifting subsystems	right prefrontal cortex
<i>Categorical-coordinate-conversion</i>	categorical spatial relation, size, and orientation of object	converts categorical relation to range of coordinates	range of coordinates to attention shifting subsystems	posterior parietal lobes
<i>Attention shifting</i>	coordinates of a perceptual unit	disengages attention from previous location, shifts to a new location, engages attention window at that location	attention window centered on perceptual unit	posterior parietal cortex, superior colliculus, thalamus

Table 2. Types of partial damage possible in the computer simulations

SUBSYSTEM	PARTIAL DAMAGE
visual buffer	there is a blind spot in the buffer
spatiotopic mapping	(a) the image is randomly mislocated (b) one location value is assigned, registering multiple objects as one
feature detection	prepares an intensity gradient of the left side of the image only
preprocessing	outputs only the trigger features, not the whole pattern
categorical relations encoding	determines taper incorrectly
coordinate relations encoding	the object center is randomly miscalculated
pattern activation	the patterns are damaged
associative memory	short-term memory is volatile
categorical property lookup	(a) repeats its first long-term memory access (b) accesses the wrong long-term memory structure
coordinate property lookup	(a) repeats its first long-term memory access (b) accesses the wrong long-term memory structure
categorical-coordinate conversion	randomly miscalculates the conversion
attention shifting	cannot replace the contents of the visual buffer in response to eye movement

Table 3. The flow of processing for normal perception when the system is asked to identify the object in a familiar fox picture. Subsystems listed in the same section are operating in parallel.

1. VISUAL BUFFER: THE ATTENTION WINDOW FOCUSES ON THE WHOLE IMAGE
2. SPATIOTOPIC MAPPING: THE OBJECT IS LOCATED IN SPACE  
PREPROCESSING: THE RELATIVELY INVARIANT TRIGGER FEATURES ARE MARKED ON THE OBJECT  
FEATURE DETECTION: A WHOLE IMAGE INTENSITY GRADIENT IS COMPUTED
3. CATEGORICAL RELATIONS ENCODING: OBJECT SIZE, LOCATION, ORIENTATION, AND TAPER CATEGORIES ARE DETERMINED  
COORDINATE RELATIONS ENCODING: OBJECT METRIC SIZE, LOCATION, CENTER, AND ORIENTATION ARE DETERMINED  
PATTERN ACTIVATION: THE INPUT TRIGGER FEATURES ARE COMPARED WITH THE TRIGGER FEATURES OF THE STORED PATTERNS, AND THEN THE INPUT PATTERN IS COMPARED WITH STORED PATTERNS
4. ASSOCIATIVE MEMORY: DORSAL AND VENTRAL INFORMATION IS ENTERED INTO ASSOCIATIVE MEMORY AND MATCHED TO STORED REPRESENTATIONS
5. CATEGORICAL PROPERTY LOOKUP: MEMORY IS ACCESSED FOR PART REPRESENTATIONS IN WHICH THE SPATIAL INFORMATION IS SPECIFIED USING CATEGORICAL RELATIONS. THE HEAD IS THE MOST DISTINCTIVE PART. THE STRENGTH OF THE ACCESSED REPRESENTATION IS RELATIVELY LOW  
COORDINATE PROPERTY LOOKUP: MEMORY IS ACCESSED FOR PART REPRESENTATIONS IN WHICH THE SPATIAL INFORMATION IS SPECIFIED IN COORDINATES. THE HEAD IS THE MOST DISTINCTIVE PART. THE STRENGTH OF THE ACCESSED REPRESENTATION IS RELATIVELY HIGH, AND HENCE THIS INFORMATION IS USED
6. ATTENTION SHIFTING: THE ATTENTION WINDOW IS SHIFTED TO THE SPECIFIED LOCATION OF THE HEAD AND SCALED TO THE SPECIFIED SIZE OF THE HEAD
7. VISUAL BUFFER: THE ATTENTION WINDOW IS FOCUSED ON A NEW LOCATION AT A NEW SCALE
8. SPATIOTOPIC MAPPING: THE SPATIOTOPIC COORDINATES OF THE PART ARE COMPUTED  
PREPROCESSING: THE RELATIVELY INVARIANT TRIGGER FEATURES ARE MARKED ON THE PART  
FEATURE DETECTION: A PART INTENSITY GRADIENT IS COMPUTED
9. CATEGORICAL RELATIONS ENCODING: PART SIZE, LOCATION, ORIENTATION, AND TAPER CATEGORIES ARE DETERMINED  
COORDINATE RELATIONS ENCODING: PART METRIC SIZE, LOCATION, CENTER, AND ORIENTATION ARE DETERMINED  
PATTERN ACTIVATION: THE INPUT TRIGGER FEATURES ARE COMPARED WITH THE TRIGGER FEATURES OF THE STORED HEAD PATTERN, AND THEN THE INPUT PATTERN IS COMPARED WITH THIS STORED PATTERN
10. ASSOCIATIVE MEMORY: DORSAL AND VENTRAL INFORMATION IS ENTERED INTO ASSOCIATIVE MEMORY AND MATCHED TO STORED REPRESENTATIONS
11. ASSOCIATIVE MEMORY: THE THRESHOLD IS REACHED; THE FOX IS IDENTIFIED

Table 4. Effects of damaging the computer simulation model in different ways. Dysfunctions are grouped according to success or failure in specific tasks. An X indicates failure in a task; no entry indicates success. Keys to the input and dysfunction abbreviations appear at the end of the table.

DYSFUNCTION GROUP	SUCCESS OR FAILURE FOR THE FIFTEEN TASKS									
	What? (foxes)				What? (faces)				Who? (faces)	
	F	UT	FO	UTO	F	R	FO	OV	F	R
									FO	OV
									F & FO	foxes
									2 F	2 F
									foxes	faces
I										
VB and AM	X	X	X	X	X	X	X	X	X	X
II										
VB-p, VB->PP, PP, PP->PA, PA, and PA->AM	X	X	X	X	X	X	X	X	X	X
III										
AM-p	X	X	X	X					X	X
IV										
VB->SM, SM-p(a), SM, SM->CoorE, CoorE, CoorE->AM, CoorL, AM->CoorL, AS, and AS->VB	X	X	X	X					X	X
V										
PP-p and PA-p	X	X	X	X					X	X
VI										
CoorE-p and CoorL-p(b)	X		X						X	X
VII										
SM->CatE, CatE-p, CatE, and CatE->AM	X	X	X	X					X	X
VIII										
CoorL->AS									X	X
IX										
CatL-p(b), CatL, AM->CatL, CatL->CCC, CoorL->CCC, CCC-p, CCC, and CCC->AS		X		X						
X										
SM-p(b)									X	X

(continued . . . .)

Table 4 (continued)

DYSFUNCTION GROUP	SUCCESS OR FAILURE FOR THE FIFTEEN TASKS															
	What? (foxes)				What? (faces)			Who? (faces)		Same?	Here?					
	F	UT	FO	UTO	F	R	FO	OV	F	R	FO	OV	F & FO	2 F	2 F	
													foxes	foxes	faces	
XI																
CoorL-p(a)				X												
XII																
CatL-p(a)					X											
XIII																
AS-p										X						
XIV																
VB->FD, FD-p, FD, and																
FD->AM																
Key to the input abbreviations:																
F	familiar				OV	overflowing				UT	unfamiliar twisted					
FO	familiar occluded				R	rotated				UTO	unfamiliar twisted occluded					
Key to the dysfunction abbreviations:																
A "-p", "-p(a)", or "-p(b)" following the initials of a subsystem denotes partial damage (see table2 for an explanation of the types of partial damage). Subsystem initials with no suffixes denote full damage, and two subsystems separated by "->" indicate a severed connection.																
AM	associative memory								CoorL	coordinate property lookup						
AS	attention shifting								FD	feature detection						
CatE	categorical relations encoding								PA	pattern activation						
CatL	categorical property lookup								PP	preprocessing						
CCC	categorical-coordinate conversion								SM	spatiotopic mapping						
CoorE	coordinate relations encoding								VB	visual buffer						

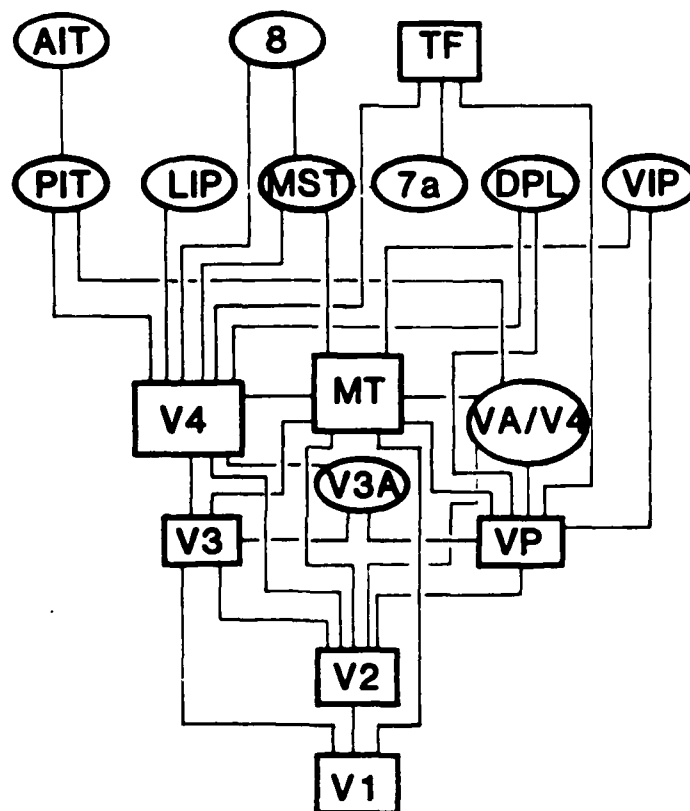
### Figure Captions

**Figure 1.** Cortical visual areas and their connections in the macaque brain. Each area is located one level above the highest level from which it receives ascending input, and is located beneath all areas from which it receives feedback (see Maunsell and Newsome, 1987; Van Essen, 1985; Van Essen and Maunsell, 1983). At higher levels of the system, areas to the left side of the figure correspond roughly to the dorsal system, and areas to the right side correspond roughly to the ventral system.

**Figure 2.** The major groups of subsystems posited by the theory. (Note that the top-down search component is not a coarse-level description of a subsystem; the subsystems that comprise it are not used only in the service of carrying out top-down search, and hence this component violates the hierarchical decomposition constraint.)

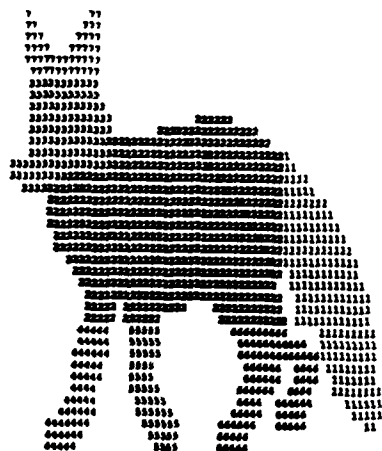
**Figure 3.** The subsystems posited by the theory.

**Figure 4.** Illustrations of four stimuli used in the computer simulations. As noted in the text, the segmentation into parts is done merely for convenience, and is not necessary in principle for the system to operate.

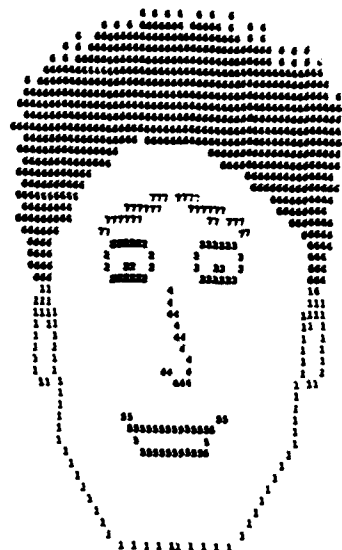




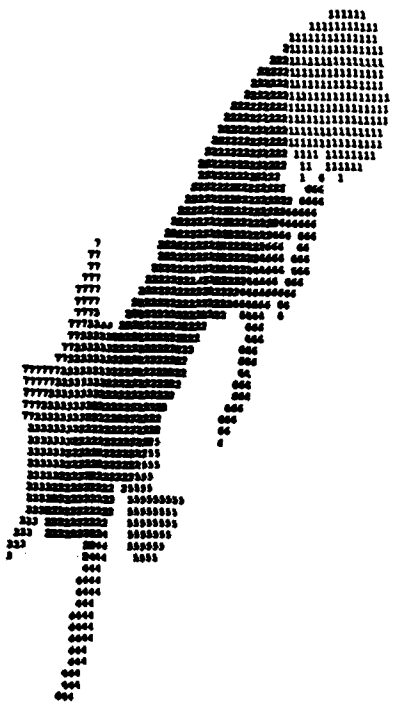




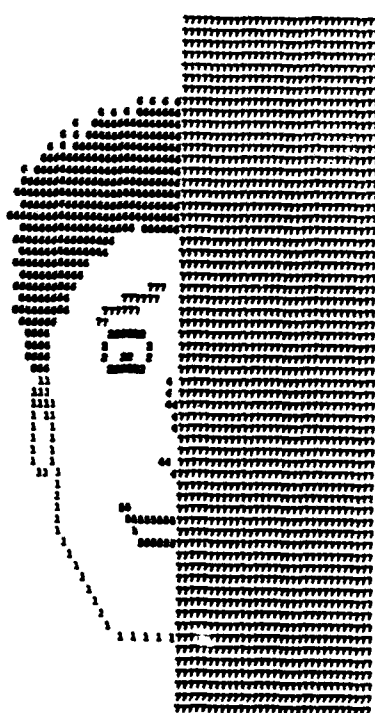
familiar fox



familiar face



unfamiliar twisted fox



familiar occluded face

Publications During Grant Period

- Kosslyn, S. M. (1988). Aspects of a cognitive neuroscience of mental imagery. *Science*, 240, 1621-1626.
- Kosslyn, S. M., Cave, C. B., Provost, D., and Von Gierke, S. (1988). Sequential processes in image generation. *Cognitive Psychology*, 20, 319-343.
- Roth, J. D., and Kosslyn, S. M. (1988). Construction of the third dimension in mental imagery. *Cognitive Psychology*, 20, 344-361.
- Arditis, A., Holtzman, J. D., and Kosslyn, S. M. (1988). Mental imagery and sensory experience in congenital blindness. *Neuropsychologia*, 26, 1-12.
- Kosslyn, S. M., Sokolov, M. A., and Chen, J. C. (1989). The lateralization of BRIAN: A computational theory and model of visual hemispheric specialization. In D. Klahr and K. Kotovsky (Eds.), *Complex Information Processing Comes of Age*. Hillsdale, NJ: Erlbaum.
- Kosslyn, S. M. (in press). Imagery. In D. Osherson, S. M. Kosslyn, and J. Hollerbach (Eds.), *An Invitation to Cognitive Science*. Cambridge, MA: MIT Press.
- Osherson, D., Kosslyn, S. M., and Hollerbach, J. (Eds.) (in press), *An Invitation to Cognitive Science*. Cambridge, MA: MIT Press.
- Kosslyn, S. M. (in press). Computational theories of imagery. *Dictionary of Cognitive Science*. London: Basil Blackwell.
- Kosslyn, S. M., and Van Kleeck, M. (in press). Broken brains and normal minds: Why humpty-dumpty needs a skeleton. In E. Schwartz (Ed.), *Computational Neuroscience*. Cambridge, MA: MIT Press.
- Kosslyn, S. M., Cave, C. B., Cronin, L. A., and Arditis, A. (in press). Visual imagery in the blind side: a neuropsychological test of the tacit knowledge hypothesis. *Brain and Cognition*
- Kosslyn, S. M. (in press). Developmental psychology. *Worldbook Encyclopedia*
- Kosslyn, S. M. (in press). Imagination. *Worldbook Encyclopedia*
- Kosslyn, S. M. (in press). Understanding charts and graphs. *Applied Cognitive Psychology*
- Kosslyn, S. M., Flynn, R. A., and Amsterdam, J. B. (in press). Components of high-level vision: A cognitive neuroscience analysis and accounts of neurological syndromes. *Cognition*
- Kosslyn, S. M., Van Kleeck, M. C., and Kirby, K. N. (in press). A neurologically plausible theory of individual differences in visual mental imagery. In J. T. E. Richardson, P. Hampson, and D. Marks (Eds.), *Advances in Mental Imagery*. London: Routledge.
- Kosslyn, S. M., Segar, C., Pani, J., and Hillger, L. A. (in press). When is imagery used? A diary study. *Journal of Mental Imagery*.
- Cave, K. R., and Kosslyn, S. M. (in press). Varieties of size-scaling in attention. *Journal of Experimental Psychology: General*.
- Kosslyn, S. M., Koenig, O., Barrett, A., Cave, C. B., Tang, J., and Gabrieli, J. D. E. (in press). Evidence for two types of spatial representations: hemispheric specialization for categorical and

coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance*

Van Kleeck, M. H., and Kosslyn, S. M. (in press). The use of computer models in the study of cerebral lateralization. In F. L. Kitterle (Ed.), *Cerebral Lateralization: Theory and Research*. Hillsdale, NJ: Erlbaum

Rueckl, J. G., Cave, K. R., and Kosslyn, S. M. (in press). Why are "what" and "where" processed by separate cortical visual systems? a computational investigation. *Journal of Cognitive Neuroscience*

Kosslyn, S. M. (in press). The psychology of visual displays. *Investigative Radiology*

#### Participating Professionals

Jay R. Rueckl, Ph.D. Assistant Professor, Department of Psychology, Harvard University (collaborator on neural network models)

John D. E. Gabrieli, Ph.D. Post doctoral fellow (departed Nov 88 to join the faculty of Northwestern University)

Olivier Koenig, Ph.D. Visiting Scholar (from the University of Geneva)

Arlette Swift, Ed.D. Post doctoral fellow (neuropsychology)

#### Advanced Degrees Awarded

J. R. Roth, Ph.D. Department of Psychology, Harvard University

In addition, six graduate students work in the laboratory, two of whom will be awarded the Ph.D. this year.

#### Presentations

Presentations were delivered at the following institutions. Unless noted otherwise, these were colloquia summarizing the material described in this Annual Report and were generally entitled "A Cognitive Neuroscience of High-Level Vision"

California Institute of Technology

University of Michigan, Ann Arbor (Business School, and Cognitive Science group; the talk at the Business School was on the psychology of visual displays)

University of Illinois, Champaign-Urbana

M.I.T.

American Association for Advancement of Science (talk was part of a symposium the PI organized)

University of North Carolina

Duke University

Association for Advancement of Artificial Intelligence (Stanford University; this talk commented on A. Newell's presentation of his SOAR universal architecture)

University of Toledo

University of Massachusetts, Amherst

Digital Equipment Corporation

McLean Hospital Brain Imaging conference

Massachusetts General Hospital (two presentations: Psychiatry residents, and Behavioral Neurology rounds)

National Institutes of Health

Children's Hospital

Boston College

University of Pennsylvania

Princeton University

Boston University

#### Consulting

National Research Council committee on Cognitive Psychophysiology

Scientific Advisory Committee, McLean Hospital

Advanced Research Initiatives review panel, Office of Naval Research (twice during period of grant).

Evaluated proposed Research Options in computer science and life science.

Consultant, Naval Research Laboratories (Stan Wilson's group). Consulted on man/machine interaction and machine representation of three-dimensional shape

James S. McDonnell Foundation Summer Institute in Cognitive Neuroscience, Director

Senior editor and co-founder: Journal of Cognitive Neuroscience

Editorial board: Psychological Review

#### Additional Progress

In addition to the work summarized above, we have just finished developing a comprehensive test battery for assessing visual mental imagery skills. The battery allows us to assess an individual's efficiency in using six component processes underlying imagery. This battery is administered by the Macintosh computer, and requires approximately 2.5 hours to complete. It currently is being given to brain-damaged subjects in order to examine the independence of the different processes; if they are indeed neurologically distinct, it should be possible to observe selective deficits in specific tasks following brain damage. In addition, the battery is being given to normal control subjects, both as a

baseline against which to compare the data from the brain-damaged subjects and in order to examine internal relations among the tasks.

The laboratory has also developed a general purpose neural network simulator, which appears to be more powerful than any simulator that is commercially available. Two versions have been implemented, one for the Macintosh II and one for a UNIX VAX environment. In addition, a program called "quick stat" has been developed to compute statistics directly on the output from our tachistoscope simulator program for the Macintosh.

Dr Jonathan Baron  
80 Glenn Avenue  
Berwyn PA 19312

Dr Jacob Beck  
Department of Psychology  
University of Oregon  
Eugene OR 97403

Dr Irving Biederman  
Dept of Psychology, Park Hall  
SUNY at Buffalo  
Amherst NY 14260

— Dr Elizabeth Bjork  
Department of Psychology  
University of California  
Los Angeles CA 90024

— Dr Robert Bjork  
Department of Psychology  
University of California  
Los Angeles CA 90024

Dr Kathryn Bock  
Department of Psychology, Uris Hall  
Cornell University  
Ithaca NY 14853

Dr L. E. Bourne, Jr  
Department of Psychology  
University of Colorado - Box 345  
Boulder CO 80309

— Dr Gordon Bower  
750 Mayfield Avenue  
Stanford CA 94305

Dr Lila Braine  
Department of Psychology  
Barnard College  
606 W 120th Street  
New York NY 10027

Dr Myron L. Braunstein  
School of Social Science  
University of California  
Irvine CA 92717

Dr Albert S. Bregman  
Department of Psychology  
1205 Dr Penfield Avenue  
Montreal Quebec  
CANADA

— Dr Doris Aaronson  
No. 29C, 110 Bleecker Street  
New York NY 10012

— Dr Jack A. Adams  
Department of Psychology  
University of Illinois  
Champaign IL 61820

Dr James A. Anderson  
Department of Psychology  
Brown University  
Providence RI 02912

— Dr John R. Anderson  
Department of Psychology  
Carnegie-Mellon University  
Pittsburgh PA 15213

Dr Nancy S. Anderson  
Department of Psychology  
University of Maryland  
College Park MD 20742

Dr Norman H. Anderson  
Psychology C-009  
University of California  
La Jolla CA 92093

Dr Fred Attneave  
Department of Psychology  
University of Oregon  
Eugene OR 97403

Dr Lloyd L. Avant  
Department of Psychology  
Iowa State University  
Ames IA 50011

Dr Harry P. Bahrick  
5 Westgate Dr  
Delaware OH 43015

Dr Donald Bamber  
Research Service  
VA Medical Center  
St Cloud MN 56301

Dr R. K. Banks  
Department of Psychology  
University of Waterloo  
Waterloo - Ontario  
CANADA

Dr Bruno Breitmeyer  
Department of Psychology  
University of Houston  
Houston TX 77004

Dr James Brennan  
Department of Psychology  
University of Mass-Harbor Campus  
Boston MA 02125

- Dr William F. Brewer  
Department of Psychology  
University of Illinois  
603 East Daniel St  
Champaign IL 61820

Dr Bruce Bridgeman  
Vision Research Lab  
University of California  
Santa Cruz CA 95064

- Dr Lee R. Brooks  
Department of Psychology  
Mc Master University  
Hamilton - Ontario  
CANADA

- Dr Nelson Butters  
3147 Morning Way  
La Jolla CA 92037

Dr Alfonso Caramazza  
Department of Psychology  
Johns Hopkins University  
Baltimore MD 21228

Dr Patricia A. Carpenter  
Department of Psychology  
Carnegie-Mellon University  
Pittsburgh PA 15213

Dr Thomas H. Carr  
Department of Psychology  
Psychology Research Building  
Michigan State University  
East Lansing MI 48824

Dr E. C. Carterette  
Department of Psychology  
University of California  
Los Angeles CA 90024

Dr Michelene Chi  
Learning Research & Development Cnt  
University of Pittsburgh  
Pittsburgh PA 15206

Dr Herbert H. Clark  
Department of Psychology  
Building 420  
Stanford University  
Stanford CA 94305

- Dr Charles Clifton  
Department of Psychology  
University of Massachusetts  
Amherst MA 01003

Dr Ronald A. Cole  
1445 Wightman St  
Pittsburgh PA 15217

Dr Lynn A. Cooper  
517 LRDC Building  
3939 O'Hara Street  
Pittsburgh PA 15206

- Dr Robert G. Crowder  
Department of Psychology  
Box 11A Yale Station  
New Haven CT 06520

Dr James E. Cutting  
Department of Psychology, Uris Hall  
Cornell University  
Ithaca NY 14853

Dr Joseph H. Danks  
Department of Psychology  
Kent State University  
Kent OH 44242

Dr James Deese  
Department of Psychology  
University of Virginia  
Charlottesville VA 22901

Dr William N. Dember  
Department of Psychology  
University of Cincinnati  
Cincinnati OH 45221

Dr Donald V. Derosa  
Department of Psychology  
Bowling Green State University  
Bowling Green OH 43403

Dr Vincent DiIollo  
Department of Psychology  
University of Alberta  
Edmonton Alta  
CANADA

Dr Emanuel Donchin  
Department of Psychology  
University of Illinois  
Champaign IL 61820

Dr Ward Edwards  
Social Science Research Inst.  
University of Southern Calif  
Los Angeles CA 90089

Dr Howard Egeth  
Department of Psychology  
Johns Hopkins University  
Baltimore MD 21218

Dr Peter D. Eimas  
Department of Psychology  
Brown University  
Providence RI 02912

— Dr Henry C. Ellis  
Department of Psychology  
University of New Mexico  
Albuquerque NM 87131

— Dr Randall Engle  
Department of Psychology  
University of South Carolina  
Columbia SC 29208

Dr William Epstein  
Psychology Building  
Charter @ W Johnson  
Madison WI 53706

Dr James R. Ertekin  
Department of Psychology  
University of Texas  
Arlington TX 76019

Dr C. W. Eriksen  
519 Psychology Building  
University of Illinois  
603 East Daniel  
Champaign IL 61820

— Dr W. K. Estes  
WM James Hall  
Harvard University  
Cambridge MA 02138

Dr Ronald Alan Finke  
Department of Psychology  
S.U.N.Y. at Stony Brook  
Stony Brook NY 11794

Dr Baruch Fischhoff  
Decision Research  
1201 Oak  
Eugene OR 97401

— Dr Ira Fischler  
Department of Psychology  
University of Florida  
Gainesville FL 32611

Dr John Flowers  
Psychology-209 Burnett  
University of Nebraska-Lincoln  
Lincoln NE 68588

Dr John M Foley  
Department of Psychology  
University of California  
Santa Barbara CA 93106

Dr Donald J. Foss  
Department of Psychology  
University of Texas  
Austin TX 78712

Dr Robert Fox  
Psychology-Wesley Hall 134  
Vanderbilt University  
Nashville TN 37240

Dr Jeffrey J. Franks  
Psychology Wesley Hall  
Vanderbilt University  
Nashville TN 37240

Dr John R Frederiksen  
Bolt Beranek & Newman  
10 Moulton Street  
Cambridge MA 02238

Dr Jennifer Freyd  
Department of Psychology  
Uris Hall  
Cornell University  
Ithaca NY 14853

Dr Wendell R Garner  
Department of Psychology  
Yale Station 11A  
New Haven CT 06520



Dr Michael S Gazzaniga  
~~Neurology~~  
 Cornell University  
 1300 York Avenue  
 New York NY 10021

Dr R. E. Geiselman  
 Department of Psychology  
 University of California  
 Los Angeles CA 90024

Dr E. Scott Geller  
 Department of Psychology  
 Virginia Tech  
 Blacksburg VA 24061

Dr Rachel Gelman  
 Department of Psychology  
 3813-15 Walnut Street  
 Philadelphia PA 19104

Dr Dedre Gentner  
 University of Illinois  
 Champaign IL 61820

Dr Eleanor J. Gibson  
 Department of Psychology  
 Uris Hall  
 Cornell University  
 Ithaca NY 14853

— Dr Murray Glanzer  
 Department of Psychology  
 New York University  
 New York NY 10003

Dr Arnold L. Glass  
 Psychology Department-Busch Campus  
 Rutgers University  
 New Brunswick NJ 08903

Dr Henry Gleitman  
 Psychology Department  
 3815 Walnut Street  
 University of Pennsylvania  
 Philadelphia PA 19104

— Dr Arthur Glenberg  
 Department of Psychology  
 University of Wisconsin  
 Madison WI 53706

Dr Sam Glucksburg  
 Department of Psychology  
 Princeton University  
 Princeton NJ 08544

Dr Walter C Gogel  
 Department of Psychology  
 University of California  
 Santa Barbara CA 93106

Dr Judith Goggin  
 Department of Psychology  
 University of Texas-El Paso  
 El Paso TX 79968

Dr Bruce Goldstein  
 Department of Psychology  
 University of Pittsburgh  
 Pittsburgh PA 15260

Dr Paula Goolkasian  
 Department of Psychology  
 University of North Carolina  
 Charlotte NC 28223

Dr Daniel Gopher  
 Department of Indus. Engin. & Mngmnt, Technion  
 Haifa 3200  
 Israel

Dr Arthur Graesser  
 Department of Psychology  
 California State University  
 Fullerton CA 92634

Dr Norma Graham  
 314 Schermerhorn Hall, Psychology  
 Columbia University  
 New York NY 10027

Dr Douglas S. Grant  
 Department of Psychology  
 The University of Alberta  
 Edmonton Alta Canada

Dr David M. Green  
 Department of Psychology  
 Harvard University  
 Cambridge MA 02138

Dr James Greeno  
 University of California at Berkley  
 Berkley CA 94720

Dr Anthony Greenwald  
 Ohio State University  
 404C W 17th Avenue  
 Columbus OH 43210

Dr Robert G. Grice  
Department of Psychology  
University of New Mexico  
Albuquerque NM 97131

Dr Steven Grossberg  
50 Hyde Street  
Newton Highland MA 02161

Dr Ralph Haber  
Psychology Department  
University of Illinois at Chicago  
Chicago IL 60680

Dr Lynn Hasher  
Department of Psychology, Weiss 8  
Temple University  
Philadelphia PA 19122

— Dr Reid Hastie  
Department of Psychology  
Northwestern University  
Evanston IL 60201

Dr Alice F. Healy  
Psychology Department  
Campus Box 345  
University of Colorado  
Boulder CO 80309

Dr Richard Held  
Department of Psychology  
79 Amherst Street  
Cambridge MA 02139

Dr Joseph B. Hellige  
Psychology Department-SEM 704  
University of Southern California  
Los Angeles CA 90089

— Dr J. V. Hinrichs  
Department of Psychology  
University of Iowa  
Iowa City IA 52242

— Dr Douglas Hintzman  
Department of Psychology  
University of Oregon  
Eugene OR 97403

Dr Ira J. Hirsh  
Department of Psychology Box 1125  
Washington University  
St. Louis MO 63130

Dr Julian Hochberg  
Department of Psychology  
Columbia University  
New York NY 10027

Dr Howard Hock  
Department of Psychology  
Florida Atlantic University  
Boca Raton FL 33431

Dr James Hoffman  
Department of Psychology  
University of Delaware  
Newark DE 19711

Dr Keith Holyoak  
Psychology Department  
University of Michigan  
330 Packard Road  
Ann Arbor MI 48104

Dr Donald Homa  
Department of Psychology  
Arizona State University  
Tempe AZ 85287

Dr Darlene V. Howard  
Department of Psychology  
Georgetown University  
Washington DC 20057

Dr James H. Howard, Jr.  
Department of Psychology  
The Catholic University  
Washington DC 20064

— Dr Earl Hunt  
Department of Psychology  
University of Washington NI-25  
Seattle WA 98195

— Dr Janellen Huttenlocher  
Department of Education  
5600 S. Kimbark Avenue  
Chicago IL 60637

Dr M.J. Peterson-Intons  
Department of Psychology  
Indiana University  
Bloomington IN 47405

— Dr Larry L. Jacoby  
Department of Psychology  
University of Utah  
Salt Lake City UT 84112

— Dr Marcia K. Johnson  
Department of Psychology  
Princeton University  
Princeton NJ 08544

Dr Neal Johnson  
Department of Psychology  
Ohio State University  
Columbus OH 43210

Dr William A. Johnston  
Department of Psychology  
University of Utah  
Salt Lake City UT 84112

Dr John Jonides  
Psychology Department  
University of Michigan  
Ann Arbor MI 48109

— Dr James Juola  
Department of Psychology  
University of Kansas  
Lawrence KS 66045

— Dr Marcel Just  
Department of Psychology  
Carnegie-Mellon University  
Pittsburgh PA 15213

Dr Daniel Kahnman  
Department of Psychology  
University of British Columbia  
Van Couver-B.C. Canada

Dr Barry Kantowitz  
Department of Psychology  
Purdue University  
West Lafayette IN 47907

Dr Stuart Katz  
Psychology Department  
University of Georgia  
Athens GA 30602

Dr Steven W. Keele  
Psychology Department  
University of Oregon  
Eugene OR 97403

Dr Frank C. Keil  
Psychology 276 Uris Hall  
Cornell University  
Ithaca NY 14853

Dr J.A. Scott Kelso  
Haskins Laboratories  
270 Crown Street  
New Haven CT 06510

— Dr Geoffrey Keppel  
Department of Psychology  
University of California  
Berkeley CA 94720

Dr Stephen M. Kerst  
35 Bayside Road #1RD  
Boston MA 02215

Dr David Kieras  
Department of Psychology  
University of Arizona  
Tucson AZ 85721

Dr John F. Kihlstrom  
Department of Psychology  
University of Wisconsin  
Madison WI 53706

— Dr Gregory A. Kimble  
Department of Psychology  
Duke University  
Durham NC 27706

Dr Ron Kinchla  
Department of Psychology  
Princeton University  
Princeton NJ 08540

Dr P. K. Kirsner  
Psychology Department  
University of Western Australia-Perth  
Nedland 6009 Australia

Dr David Klahr  
Department of Psychology  
Carnegie Mellon University  
Pittsburgh PA 15213

— Dr Stuart T. Klapp  
Department of Psychology  
California State University  
Hayward CA 94542

— Dr Roberta L. Klatzky  
Psychology Department  
University of California, Santa Barbara  
Santa Barbara CA 93106

Dr Sylvan Kornblum  
Mental Health Research Institute  
University of Michigan  
Ann Arbor MI 48109

Dr Stephen M. Kosslyn  
Psychology & Social Relations  
Harvard University  
Cambridge MA 02138

Dr Judith Kroll  
Department of Psychology & Education  
Mount Holyoke College  
South Hadley MA 01075

Dr Lester E. Krueger  
Human Performance Center  
404B W. 17th Avenue  
Columbus OH 43210

Dr Carol L. Krumhansl  
Department of Psychology  
Uris Hall  
Cornell University  
Ithaca NY 14853

Dr Michael Kubovy  
Psychology Department  
Tillett Hall  
Rutgers University  
New Brunswick NJ 08903

Dr David Laberge  
School of Social Science  
University of California  
Irvine CA 92717

Dr Marcy Lansman  
Psychology Department Davie Hall 013A  
University of North Carolina  
Chapel Hill NC 27514

Dr Joseph L. Loftus  
Department of Psychology  
Vanderbilt University  
Nashville TN 37240

Dr Lester Lefton  
Department of Psychology  
University of South Carolina  
Columbia SC 29208

Dr Alvin Liberman  
Haskins Laboratories  
270 Crown Street  
New Haven CT 06510

Dr Gregory Lockhead  
Department of Psychology  
Duke University  
Durham NC 27706

- Dr Elizabeth Loftus  
Department of Psychology  
University of Washington  
Seattle WA 98195

Dr Gregory Loftus  
Department of Psychology  
University of Washington  
Seattle WA 98195

Dr Gordon D. Logan  
Department of Psychology  
University of Illinois  
Champaign IL 61820

Dr Jack M. Loomis  
Department of Psychology  
University of California  
Santa Barbara CA 93106

Dr Lola Lopes  
Department of Psychology  
University of Wisconsin  
Madison WI 53706

Dr R. Duncan Luce  
William James Hall  
Harvard University  
Cambridge MA 02138

- Dr George Mandler  
Center Human Information Processing  
University of California  
La Jolla CA 92093

- Dr Jean M. Mandler  
Department of Psychology  
University of California  
La Jolla CA 92093

Dr Ellen M. Markman  
Psychology Building 420  
Stanford University  
Stanford CA 94305

Dr Dominic Massaro  
Program In Experimental Psychology  
University of California  
Santa Cruz CA 95064

Dr James L. McClelland  
Department of Psychology  
Carnegie Mellon University  
Pittsburgh PA 15213

Dr George McConkie  
Ctr For Study of Reading  
51 Gerty Drive  
Champaign IL 61820

— Dr Gail McKoon  
Psychology Department  
Northwestern University  
Evanston IL 60201

— Dr Douglas C. Medin  
Department of Psychology  
University of Illinois  
Champaign IL 61820

Dr David E. Meyer  
1715 Shadford Rd  
Ann Arbor MI 48104

Dr Glenn Meyer  
Department of Psychology  
Lewis and Clark College  
Portland OR 97219

— Dr George A. Miller  
Psychology Department  
Green Hall  
Princeton University  
Princeton NJ 08544

Dr Jeffrey Miller  
Department of Psychology C-009  
University of California  
La Jolla CA 92093

Dr Morris Moscovitch  
Department of Psychology  
Erindale College  
Mississauga Ont Canada

Dr Jerome L. Meyers  
Department of Psychology  
University of Massachusetts  
Amherst MA 01003

Dr Jacob Nachmias  
Department of Psychology  
3815 Walnut Street  
Philadelphia PA 19104

Dr David Navon  
Department of Psychology  
University of Haifa  
Haifa-31999 Israel

— Dr James H. Neely  
Psychology Department  
Purdue University  
West Lafayette IN 47907

— Dr Ulric Neisser  
Psychology Department  
Emory University  
Atlanta GA 30322

— Dr Thomas O. Nelson  
Psychology Department (NI-25)  
University of Washington  
Seattle WA 98195

Dr Raymond S. Nickerson  
Bolt Beranek and Newman  
50 Moulton Street  
Cambridge MA 02138

— Dr Donald Norman  
Department of Psychology  
University of California  
La Jolla CA 92093

Dr Gregg Oden  
Department of Psychology  
University of Wisconsin  
Madison WI 53706

Dr Judith Reitman Olson  
Comp. Infor. Systems-Bus. Grad. Sch  
University of Michigan  
Ann Arbor MI 48109

Dr Gary Olson  
University of Michigan  
Perry Building  
330 Packard Road  
Ann Arbor MI 48104

Dr Andrew Ortony  
174 Reading Ctr.  
51 Gerty Drive  
Champaign IL 61820

Dr Kenneth R. Paap  
Department of Psychology  
New Mexico State University  
Las Cruces NM 88003

Dr Raja Parasuraman  
Human Performance Lab  
Catholic University of America  
Washington DC 20064

Dr James Pellegrino  
Grad School of Education  
University of California  
Santa Barbara CA 93106

Dr Charles A. Perfetti  
LRDC  
University of Pittsburgh  
Pittsburgh PA 15260

Dr Kathy Pezdek  
Department of Psychology  
Claremont Graduate School  
Claremont CA 91711

Dr Herbert L. Pick, Jr  
Inst of Child Development  
University of Minnesota  
Minneapolis MN 55455

Dr Steven Pinker  
Department of Psychology  
E-10-018  
Massachusetts Institute of Technology  
Cambridge MA 02139

Dr David b. Pisoni  
Department of Psychology  
Indiana University  
Bloomington IN 47405

Dr R. H. Pollack  
Department of Psychology  
University of Georgia  
Athens GA 30602

Dr Irwin Pollack  
Mental Health Resch Institute  
University of Michigan Stop 053  
Ann Arbor MI 48104

- Dr Alexander Pollatsek  
Department of Psychology  
University of Massachusetts  
Amherst MA 01003

Dr James R. Pomerantz  
Department of Psychology  
4230 Ridge Lea Road  
Buffalo NY 14226

Dr Michael I. Posner  
Department of Psychology  
University of Oregon  
Eugene OR 97403

- Dr L. J. Postman  
Department of Psychology  
University of California  
Berkeley CA 94720

- Dr Mary Potter  
Department of Psychology  
E10-032  
Massachusetts Institute of Technology  
Cambridge MA 02139

Dr George R. Potts  
Department of Psychology  
University of Denver  
Denver CO 80208

Dr Robert Proctor  
Department of Psychology  
Auburn University  
Auburn AL 36849

Dr Dennis R. Proffitt  
Psychology Department  
Gilmer Hall  
University of Virginia  
Charlottesville VA 22901

Dr Dean G. Purcell  
Department of Psychology  
Oakland University  
Rochester MI 48063

Dr David H. Raab  
Department of Psychology  
Brooklyn College  
Brooklyn NY 11210

Dr Jan Rabinowitz  
Department of Psychology  
Barnard College  
606 W. 120th Street  
New York NY 10027

- Dr Roger Ratcliff  
Psychology Department  
Northwestern University  
Evanston IL 60201

Dr Keith Rayner  
Department of Psychology  
Tobin Hall  
University of Massachusetts  
Amherst MA 01003

Dr Arthur S. Reber  
Department of Psychology  
Brooklyn College  
Brooklyn NY 11210

— Dr Lynne Reder  
Department of Psychology  
Carnegie-Mellon University  
Pittsburgh PA 15260

— Dr Stephen Reed  
Psychology Department  
Florida Atlantic University  
Boca Raton FL 33431

Dr Robert Rescorla  
Department of Psychology  
University of Pennsylvania  
3815 Walnut Street  
Philadelphia PA 19104

Dr Lance Rips  
Behavioral Sciences  
5848 S. University Avenue  
Chicago IL 60637

Dr Irvin Rock  
Psychology Department-Livingston  
Rutger University  
New Brunswick NJ 08903

— Dr Henry L. Roediger, III  
Department of Psychology  
Purdue University  
West Lafayette IN 47907

Dr Eleanor R. Ruch  
Department of Psychology  
University of California  
Berkeley CA 94720

Dr David A. Rosenbaum  
School of Cognitive Sci Comm.  
Hampshire College  
Amherst MA 01002

— Dr David C. Rubin  
Department of Psychology  
Duke University  
Durham NC 27706

— Dr Dewey Rundus  
Department of Psychology  
University of South Florida  
Tampa FL 33620

— Dr Timothy A. Salthouse  
Department of Psychology  
University of Missouri  
Columbia MO 65211

Dr James Sawusch  
Department of Psychology  
State University of New York  
4230 Ridge Lea Road  
Buffalo NY 14226

Dr Arthur Schulman  
Department of Psychology  
University of Virginia  
Charlottesville VA 22901

Dr Roger Schvaneveldt  
Department of Psychology  
New Mexico State University  
Las Cruces NM 88003

Dr Barry Schwartz  
Psychology Department  
Swarthmore College  
Swarthmore PA 19081

Dr Richard Schweickert  
Psychological Sciences  
Purdue University  
West Lafayette IN 47907

Dr J. G. Seamon  
Department of Psychology  
Wesleyan University  
Middletown CT 06457

Dr Robert Sekuler  
2003 Orrington Avenue  
Evanston IL 60201

Dr James Shanteau  
Department of Psychology  
Kansas State University  
Manhattan KS 66506

Dr Marilyn Shatz  
Human Performance Center  
330 Packard Road  
Ann Arbor MI 48104

Dr Wayne Shebilske  
Comm on Vision  
National Academy of Science  
2101 Constitution A  
Washington DC 20418

Dr Roger N. Shepard  
Psychology Department  
Building 420  
Stanford University  
Stanford CA 94305

— Dr Richard M. Shiffrin  
Department of Psychology  
Indiana University  
Bloomington IN 47405

— Dr Edward J. Shoben  
Psychology Department  
University of Illinois  
603 East Daniel Street  
Champaign IL 61820

— Dr Harvey Shulman  
Human Performance Center  
404-B West 17th Ave.  
Columbus OH 43210

Dr Robert Siegler  
Department of Psychology  
Carnegie-Mellon University  
Pittsburgh PA 15260

Dr H. A. Simon  
Department of Psychology  
Carnegie-Mellon University  
Pittsburgh PA 15260

— Dr Norman J. Slamecka  
Department of Psychology  
University of Toronto  
Toronto Ont M5S1A1 Canada

Dr Edward E. Smith  
Bolt Beranek & Newman Inc  
50 Moulton Street  
Cambridge MA 02139

Dr Linda B. Smith  
Department of Psychology  
Indiana University  
Bloomington IN 47405

Dr Robert Solso  
Psychology Department  
University of Nevada, Reno  
Reno NV 89557

Dr Robert Sorkin  
Department of Psychology  
Purdue University  
West Lafayette IN 47907

Dr George Sperling  
Psychology Department  
New York University  
6 Washington Place RM 980  
New York NY 10003

— Dr Kathryn T. Spoehr  
Psychology Department  
Brown University  
Providence RI 02912

Dr Larry Squire  
Vet. Adm. Medical Center  
3350 La Jolla Village Drive  
San Diego CA 92161

Dr Keith Stanovich  
Department of Psychology  
Oakland University  
Rochester MI 48063

Dr Robert Sternberg  
Department of Psychology  
Box 11A Yale Station  
New Haven CT 06520

Dr Saul Sternberg  
Department of Psychology  
University of Pennsylvania  
Philadelphia PA 19104

Dr John Swets  
C/O Bolt Beranek & Newman  
50 Moulton Street  
Cambridge MA 02138

Dr David A. Swinney  
Psychology Department  
Tufts University  
Medford MA 02155

Dr Michael Tanehaus  
Psychology Department  
University of Rochester  
Rochester NY 14726



— Dr John Theios  
Department of Psychology  
University of Wisconsin  
Madison WI 53706

Dr James T. Townsend  
Department of Psychology  
Purdue University  
West Lafayette IN 47907

Dr Tom Trabasso  
Department of Psychology  
University of California  
Santa Barbara CA 93106

Dr Anne M. Treisman  
Psychology Department  
University of British Columbia  
Vancouver BC Canada

Dr Barbara Tversky  
Department of Psychology  
Building 420  
Stanford University  
Stanford Ca 94305

Dr Ovid Tzeng  
Department of Psychology  
University of California  
Riverside CA 92521

Dr James Voss  
LRDC-University of Pittsburgh  
3939 Ohara Street  
Pittsburgh PA 15260

Dr Brian A. Wandell  
Department of Psychology  
Stanford University Bldg. 420  
Stanford CA 94305

Dr J. G. Snodgrass  
New York University  
6 Washington Place Rm 856  
New York NY 10003

Dr Robert Weber  
Department of Psychology  
Oklahoma State University  
Stillwater OK 74078

Dr Daniel Weintraub  
Human Performance Center  
Perry Building  
330 Packard Road  
Ann Arbor MI 48104

Dr Wayne Wickelgren  
Department of Psychology  
University of Oregon  
Eugene OR 97403

Dr Delos D. Wickens  
Psychology Department  
Colorado State University  
Fort Collins CO 80525

Dr Thomas D. Wickens  
Department of Psychology  
University of California  
Los Angeles CA 90024

Dr George Wolford  
Department of Psychology  
Dartmouth College  
Hanover NH 03755

Dr Keith Wollen  
Department of Psychology  
Washington State University  
Pullman Wa 99164

Dr Franks Yates  
Department of Psychology  
University of Michigan  
Ann Arbor MI 48104

— Dr Rose Zacks  
Department of Psychology  
Michigan State University  
East Lansing MI 48824

— Dr Eugene B. Zechmeister  
Department of Psychology  
6252 N. Sheridan Road  
Chicago IL 60626